# DOCUMENT RESUME

ED 201 473 SE 034 611

| | |
|---|---|
| AUTHOR | Singer, Arlene |
| TITLE | Option Y, Statistics. |
| INSTITUTION | Hawaii State Dept. of Education, Honolulu. Office of Instructional Services. |
| REPORT NO | RS-80-9834 |
| PUB DATE | Sep 80 |
| NOTE | 96p. |
| EDRS PRICE | MF01/PC04 Plus Postage. |
| DESCRIPTORS | Annotated Bibliographies; *Cognitive Objectives; *Course Objectives; Instructional Materials; Mathematical Applications; *Mathematics Instruction; Problem Solving; *Resource Materials; Secondary Education; *Secondary School Mathematics; *Statistics; Teaching Methods |


ABSTRACT

This guide outlines a one semester Option Y course, which has seven learner objectives. The course is designed to provide students with an introduction to the concerns and methods of statistics, and to equip them to deal with the many statistical matters of importance to society. Topics covered include graphs and charts, collection and organization of data, measures of central tendency and dispersion, uses and misuses of statistics, frequency distributions, correlations and regression. The document opens with sections on the nature of statistics, its value, problem solving, approaches to teaching, and the role and nature of audio-visual aids. The guide has a section on each of the seven objectives, which lists important concepts, performance objectives, sample exercises, and suggested student reading and film/slide presentations. The guide concludes with a comprehensive list of available materials, including: (1) annotated bibliographies of texts and supplementary books; (2) additional resources; (3) sources of data; (4) films; (5) slides/cassettes; (6) transparencies; (7) miscellaneous; (8) publishers; and (9) sources of audiovisual materials. (MP)

*"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write."*

*H.G. Wells*

## FOREWORD

In 1978 the Mathematics Program Guide, K-12 was developed and disseminated to all public schools in Hawaii "to provide direction for teachers and administrators in the development of school-level mathematics." One of the major outcomes of this effort was a substantial strengthening of the quantity and quality of the courses offered as part of the secondary mathematics program. Existing courses in grades 9-12 were restructured and several new courses were created.

This document is a course guide for the development of a one-semester statistics course. It provides teachers with guidelines, suggested activities, and a resource list to use in structuring the course. The guide recommends, rather than limits the content of the course. Teachers should be stimulated to expand upon activities as they work toward providing students with experiences designed to meet each specified objective.

The Office of Instructional Services gratefully acknowledges those teachers who critically evaluated the draft manuscript of this guide. Special recognition is extended to Arlene Singer, who developed this course guide.

Charles G. Clark
Superintendent of Education

## TABLE OF CONTENTS

INTRODUCTION

# THE NATURE OF STATISTICS

"Statistics is the art of making numerical conjectures about puzzling questions." (1)

"Statistics ... has been defined as a science where one studies how to make wise decisions in the face of uncertainty." (5)

"Statistics [is] the science concerned with collection, analysis, and interpretation of numerical information ... " (2)

As used in the above quotations, the word statistics describes a field of mathematics. However, the word statistics can also refer to single pieces of numerical data. Statistical methods are used to study and analyze numerical data about specific events.

When it is practical to gather all the pertinent data in solving a problem, a branch called descriptive statistics is used. Descriptive statistics describes or summarizes data in a clear and precise manner. No guesswork is involved and no inferences about the data are necessary; all data can be gathered by counting or measuring.

- How many students at University High School earned a grade of A in mathematics last term?
- What is the mean number of books circulated each month at Hamilton Library?
- What was the median salary of the sales staff at Akamai Sales Co. last month?

The data necessary to accurately answer each of the above questions is available; thus, descriptive statistics would be used.

In contrast to this is another branch of statistics called inferential statistics. Inferential statistics is used when it is not practical or possible to gather all the data necessary to answer a question or solve a problem. Information is collected for a carefully selected sample of the entire set, or population. The methods of inferential statistics are then used to make generalizations and predictions about the population. Answering the following questions requires inferential statistics:

- What was Hawaii's unemployment rate during the month of April?
- In a carload of lightbulbs, what percent of bulbs will burn out in less than 1000 hours?
- What was Hawaii's most watched television show during the week of October 7?

Nonstatisticians usually find descriptive statistics more useful and easier to understand than inferential statistics. Facts are easier to deal with than probabilities and uncertainties. Researchers and scientists, however, find inferential statistics to be a powerful tool. It allows them to intelligently analyze experimental data, which is frequently incomplete.

Statistics in one form or another is used in virtually every field to answer questions or to aid in making the decisions which are most likely to be correct. It is used to determine market trends, to monitor production quality, to keep track of public opinion, to set standards, to determine safety levels, and so on.

Statistical methods are powerful tools used to study the past and the present as well as to extrapolate to the future. Statistics also plays a large role in the science or art of decision making.

References


1. Freedman, David, Robert Pisani and Roger Purves, Statistics, W. W. Norton & Co., Inc., N.Y., 1978.

2. Overview and Analysis of School Mathematics, K-12, Conference Board of the Mathematical Sciences, Washington, D.C., 1975 (NACOME Report).

3. Pieters, Richard S. and John J. Kinsella, "Statistics," The Growth of Mathematical Ideas, Grades K-12, 24th Yearbook, NCTM, 1959.

4. Schmidt, Marty J., Understanding and Using Statistics, Basic Concepts, D. C. Heath and Company, 1975.

5. Synopses for Modern Secondary School Mathematics, Organization for Economic Cooperation and Development, 1966.


9

## THE VALUE OF STATISTICS

*Statistics, the science concerned with collection, analysis, and interpretation of numerical information is important in the life of every citizen. It is needed for the proper evaluation of everyday matters such as advertising claims about gasoline mileage and relief from indigestion, public opinion polls and weather reports. It is indispensable for the solution of policy questions, from local affairs such as property assessment and predictions of school enrollments to national problems involving unemployment, crime, airplane safety and health. Even though numerical information is encountered everywhere, in newspapers and in magazines, on radio and on television, too few people have the training to accept such information critically and use it effectively.*

<u>Overview and Analysis of School Mathematics, Grades K-12, 1975</u>, Conference Board of the Mathematical Sciences, Washington, D.C. 1975

No matter which way people today turn, they cannot avoid statistical statements such as:

- 9 out of 10 doctors recommend ...
- The cost of living rose by 6% this year ...
- There is an 80% chance of rain ...
- The mean income for a family of four ...
- The unemployment rate dropped 0.2% last month ...

Statistics can be and are frequently used to inform, convince, cajole, entertain, and mislead us. People need to be able to separate fact from fiction and half-truths. They must realize that a cause and effect relationship does not necessarily exist whenever a correlation between two things occurs. Any information provided must be analyzed carefully before an intelligent decision can be reached about its value. People who have learned to do this analysis, either consciously or unconsciously, are able to make better informed decisions about things which affect their lives.

Statistics is an essential part of modern science and technology. Demand for statistical decision-making continues to rise in these areas and many of these decisions have a direct influence on our lives. For example, any time a food or drug item is found hazardous to our health and removed from the market, statistics has played a major role in the decision. It is possible that carefully controlled experiments have shown a very high correlation between the item in question and susceptibility to disease.

Obviously, a one-semester statistics course will not prepare students to conduct a comprehensive survey. However, at the end of the course, students should have an appreciation for what statistics can and cannot do. They should also have an understanding of basic statistical reasoning to aid them in everyday life. Furthermore, many students will find the need to use statistics as a tool in some future profession. And every one of these students will at some time find the need to take a stand on an issue based on statistical reasoning. Since many decisions must be made in the face of uncertainty, it is also hoped that the study of statistics will enable the students to make these decisions with a little more confidence.

As early as 1959 at least two major groups recognized the need for statistics in the school curriculum. The Commission on Mathematics of the College Entrance Examination Board suggested that statistics be among the new content of school instruction. In that same year the National Council of Teachers of Mathematics (NCTM) published a yearbook entitled The Growth of Mathematical Ideas, Grades K-12. One entire chapter was devoted to the teaching of statistics. In addition to providing examples, the authors indicate when and how statistical concepts should be introduced in the classroom.

In 1967 the Joint Committee on the Curriculum in Statistics and Probability was formed by NCTM and the American Statistical Association (ASA). These two groups felt that broad based efforts were necessary if statistics is to take its place in society. They focused on two tasks: first, persuading educators and parents that statistics is important and that it should become part of the curriculum and, second, preparing materials which teachers could utilize and consult (see annotated listing of supplementary materials).

In 1975 the NACOME report was published by the Conference Board of the Mathematical Sciences. It recommended that "instructional units dealing with statistical ideas be fitted throughout the elementary and secondary school curriculum." One possible course is described as:

> "A statistics course for high school students with little or no algebra, especially for non-college bound students in the social sciences who as consumers and citizens must learn to cope with numerical information. The main theme of such a course would be 'making sense out of numbers' without getting involved with complicated mathematical formulae."

It also recommended the emphasis of statistics "as an interdisciplinary subject with applications in the natural, physical and social sciences and the humanities."

The NCTM annual meeting in April, 1981 will also emphasize the teaching of statistics. Their yearbook entitled Statistics and Probability will be presented at this meeting. Articles will fall into three categories: background, teaching of statistics and probability, and using statistics. The focus will be on practical rather than theoretical considerations.

The use of statistical methods is a powerful tool which, whether applied with mathematical rigor or just common sense, affects the lives of everyone. The study of statistics not only provides students with the chance to apply many of the mathematical skills learned, but also improves critical thinking and problem-solving abilities. It can also be used "to emphasize important processes such as active inquiry, discovery of relationships, the testing of conjectures, and critical thinking" (8).

References

1. Campbell, Stephen K., Flaws and Fallacies in Statistical Thinking, Prentice-Hall, Inc., 1974.

2. Klitz, Ralph H. and Joseph F. Hofmeister, "Statistics in the Middle School," Arithmetic Teacher, NCTM, February 1979, p. 35-36, 59.

3. Overview and Analysis of School Mathematics, K-12, Conference Board of the Mathematical Sciences, Washington, D.C., 1975 (NACOME report).

4. Pieters, Richard S. and John J. Kinsella, "Statistics," The Growth of Mathematical Ideas, Grades K-12, 24th yearbook, NCTM, 1959.

5. Råde, Lennart, editor, The Teaching of Probability and Statistics, Proceedings of the first CSMP International Conference, John Wiley and Sons, 1970.

6. Schmidt, Marty J., Understanding and Using Statistics, Basic Concepts, D. C. Heath and Co., 1975.

7. Shulte, Albert P., "A Case for Statistics," Arithmetic Teacher, NCTM, February 1979, p. 24.

8. Statistics and Information Organization, Mathematics Resource Project, University of Oregon, Creative Publications, 1978.

# PROBLEM SOLVING AND STATISTICS

"... when the need to use statistical methods arises, it is always because there is a problem to be solved, a decision to be made, or a question to be answered." (1)

It is impossible to study statistics without encountering problem solving techniques and strategies. One of these techniques, for example, the use of graphs and tables, is especially helpful in organizing data and looking for trends and patterns.

George Polya, a twentieth-century mathematician, has identified four steps to problem solving: understanding the problem, devising a plan, carrying out the plan, and looking back (see pages 15-23 in the State's Mathematics Program Guide). These four steps are just as applicable to the solution of statistical problems as they are to problems in any other field. Often the steps are intertwined. Reliable conclusions are the result of good problem solving involving clear thinking, careful work, and often some ingenuity.

Step 1: Before any problem can be solved it must be understood. One of the more difficult things a mathematician must do is to define the problem carefully and unambiguously. All key words must be defined accurately and all assumptions and relationships must be recognized. For example, in the question "How many students in Mrs. Horiuchi's math class received a B?" should we be concerned with students who received B+ or B-? To which one of Mrs. Horiuchi's math classes does the question refer? Or does it refer to all of her classes? Are we interested in knowing how many B's were received on the last test? In the last quarter? Semester? Or year?

Other important aspects of understanding a problem include identifying and discarding irrelevant facts and deciding how exact the solution should be. Problem solving behaviors associated with this step include telling the problem to someone else, stating the conditions of the problem, making a table or a graph, outlining the problem, and using resources to clarify terms.

Step 2: Devising a plan, includes identifying what additional information is needed, where it can be found, and how it can best be collected. Can descriptive statistics be used, i.e., are all of the relevant data available? Or should a sample be selected and inferential statistics be applied? This is often the most difficult stage in problem solving. Teachers can help students by modeling the technique, providing students with a checklist of choices, or asking questions that suggest choices.

Step 3: Carrying out the plan includes collecting the data, organizing and analyzing the data to the accuracy necessary for solution of the problem. Students should be encouraged to keep the problem continuously in mind. Also, in collecting the data, students may find that their plan needs to be revised.

Step 4: In looking back at the problem there are many questions that must be considered. It is important to check both the results and the process used. Questions and concerns that arose during the earlier stages of problem solving must also be reflected upon. If any assumptions were identified in

Step 1, the solution must take them into consideration. Were all of the discarded facts really "irrelevant?" Has the problem been adequately solved? Is the answer reasonable? If a sample was chosen, is it sufficiently large and unbiased? Is it a good sample from which to make a prediction? If a prediction has been made, how much confidence can be placed in it? Are any of the graphs misleading? Are there any other factors which would influence the solution?

Problem solving and statistical reasoning are interrelated in real life. Not only are problem-solving techniques necessary for statistics research, but also statistical methods are often used as an aid in solving other types of problems. To answer the question, "Should a measles/mumps vaccine be required for all students enrolling in school in Hawaii?" statistical methods can be applied. The appropriate data would be collected, organized, and analyzed. Correlations between variables might be found and predictions would be made. Based on this analysis a decision could be made. The solution might not be exact or certain, but it could be suggested with a measurable amount of confidence. As concluded by the Mathematics Resource Project at the University of Oregon, "Students must learn to use statistics ... in problem solving to give reasonable answers when faced with uncertain stituations."

References

1. Armore, Sidney J., Introduction to Statistical Analysis and Inference, John Wiley & Sons, Inc., 1966.

2. "Collecting, Organizing, and Interpreting Data," Booklet 16, More Topics in Mathematics, 30th Yearbook, NCTM, 1969.

3. Pieters, Richard S. and John J. Kinsella, "Statistics," The Growth of Mathematical Ideas, K-12, 24th Yearbook, NCTM, 1959.

4. Statistics and Information Organization, Mathematics Resource Project, University of Oregon, Creative Publications, 1978.

## APPROACHES TO TEACHING STATISTICS

Too often students perceive math as a class in which homework is discussed, a new method or procedure for solving a problem is demonstrated, and homework is assigned. In general they consider mathematics boring. Mathematics, however, can and should be an exciting subject.

Over the years, teachers have developed many approaches to teaching mathematics that, if used properly, offer varied and often exciting insights to the topic being studied. Some of these methods are described below. None of them are mutually exclusive; many excellent lessons involve a combination of these and other approaches.

1. *Audio-visual presentations:* Audio-visual materials provide a versatile way to explore many topics. Films, cassettes, slides, and transparencies are readily available and often extremely helpful. (See the Audio Visual Materials section of this guide for specific examples.)

2. *Discovery Lessons:* Through careful organization or the appropriate timing of a question, a teacher can often *guide* students into discovering "new" concepts. For example, a well-planned exercise can lead a student to discover that the sum of the deviations of the data from the mean is always zero. $(\Sigma(x - \bar{x}) = 0)$

3. *Laboratory Lessons:* Laboratory lessons emphasize learning by doing and are often a type of discovery lesson. The focus is on the learning process rather than the teaching process. Many of these lessons involve experimentation and the collection, organization and analysis of data. One example is to have students prepare and conduct a survey on an issue relevant to the student body. Students will have to consider questions such as: What is the purpose of the study? What questions must be asked to achieve this purpose? Will the questions result in meaningful data? How can the information obtained be best summarized and interpreted?

4. *Lessons Evaluating Sources of Data:* Students should realize that there is often reason to doubt the validity of the data presented. Knowledge of the source of data often raises questions concerning possible biases in the collection, presentation, and analysis of it. Guidelines for evaluating these sources might include the following questions: Who is the authority? What is his/her background? What is the purpose of the material? Is the authority objective in the treatment of the material?

5. *Open Textbook Studies:* More often than not students do not read their textbooks because they do not know how. As a result the only part of the text a student is familiar with is the pages of exercises. If he/she has difficulty with a problem, it just doesn't get done. The text is not consulted for hints or explanations or even definitions. There should be periods when the students *read* and *explore* their course materials with their teacher. This should help to teach students the proper way to read a math text--it can't be read like a novel.

6. *Use of Cartoons:* Sometimes cartoons help students fix concepts in their minds. For example, a student may remember to question the use of the word average if a cartoon has been shown of a man submerged in a pool of water and the caption: average depth = 3 feet.

7. *Use of Calculators/Computers:* The availability of calculators and the increasing availability of computers in schools today can provide valuable classroom learning aids. Until recently most problems and exercises were chosen on the basis of ease of computation. Now more realistic examples are possible without undue emphasis on the mechanics of the problem.

8. *Use of Models:* Concrete models can be used to study abstract mathematical ideas and mathematical models often help in the study of concrete real-life situations. Graphs are models in statistics which convey information very quickly to the eye. Scatter diagrams model possible relationships between two quantities.

9. *Use of Newspapers and Magazines:* Articles clipped by the teacher or brought in by the student can be used to demonstrate real-world applications of statistics. They can also be analyzed to determine whether the data has been presented properly or whether it is in fact misleading.

10. *Use of Television:* Students can keep a log of commercials that use statistics to impress the consumer. Also, news programs use statistics regularly. They can be classified as misleading or acceptable with reasons cited. Are the sources of the material identified? Were all the facts presented? Does the conclusion seem justified on the basis of the facts presented?

## THE ROLE OF AUDIO-VISUAL AIDS IN TEACHING STATISTICS

The use of audio-visual aids in the classroom is essential. They make mathematics more meaningful to students by adding variety, depth and breadth to the learning process. The students' number sense, space perception and imagination necessary to master mathematical ideas can be guided and built upon by teachers and, when used properly, these materials can help. Also, using audio-visual materials should result in more correct and richer learning, more efficient use of time and improve student retention.

The teacher should understand, however, that different devices offer unique contributions and that not every aid is suited for every purpose. If they are not used effectively, these materials can be a waste of both time and money.

Different types of audio and visual materials are used to attain the following goals, adapted from the book Guidelines for Teaching Mathematics by Donavan A. Johnson and Gerald A. Rising:

1. *Visualize abstract ideas so they have meaning.* For example, the film "Fact Finder for the Nation" by the U.S. Census Bureau will enrich the students' understanding of the terms statistics and survey among others.

2. *Illustrate applications of mathematics in our world.* The newspapers are full of examples of how statistics is used in our world.

3. *Bring to the school and the classroom important first-hand accounts of new activities in mathematics and mathematics education.* A tape recorded session from an NCTM meeting on the relevance of statistics in today's society can be used as an introduction to the course.

4. *Build favorable attitudes toward and interest in mathematics.*

5. *Present the history of mathematics* (statistics) *and other enrichment topics.*

6. *Illustrate the discovery of relationships or principles.* There are some excellent transparency sets available for this purpose.

7. *Present dynamic ideas that depend on motion.*

8. *Correlate mathematics with other subjects by presenting supplementary materials.* The film "Using SQC" (Statistical Quality Control) produced by the National Food Processors Association shows how statistics is used in jobs concerned with quality control.

9. *Provide complex drawings for 3-dimensional effects.*

10. *Teach how to solve problems.* Again, many transparency sets are available.

17

11. *Introduce a new subject or unit.* Correlation can be introduced, for example, with the cartoon of a person sitting in a bathtub with the telephone sitting on a table right next to the tub. The caption: This time I'm ready for you to ring!

12. *Summarize or review units within a course.* Review and summary may be done with a short film on statistics.

# THE COURSE

As stated in the Mathematics Program Guide, Statistics, an Option y course, is intended "to provide students with an introduction to the concepts and methods of statistics. It will equip them to deal with the many statistical matters of importance to society. Topics included in this course are graphs and charts, collection and organization of data, measures of central tendency and dispersion, uses and misuses of statistics, frequency distributions, correlation and regression."

There are seven learner objectives for this course. These objectives indicate the minimum course content for a Statistics course. The student:

I. Distinguishes between descriptive and inferential statistics.

II. Constructs frequency distributions and displays the data using various graphical methods.

III. Computes percentile and percentile ranks.

IV. Calculates and compares the mean, the median, and the mode.

V. Describes how scores are dispersed about the center by computing range, average deviation, variance, and standard deviation.

VI. Calculates linear correlation coefficients and draws scattergrams.

VII. Predicts one variable's value, knowing another, by using the least squares regression equation and line.

The remaining pages of this section address each of the objectives individually. Information relative to each objective is provided under the headings: Important Concepts, Performance Objectives of the Course, Sample Exercises and Solutions, and Suggested Student Reading and Film/Slide Presentations.

Important Concepts includes new vocabulary and definitions. Notation and formulas are given where appropriate and examples of many concepts are provided. A listing of performance objectives (as used in this document refers to identified course objectives for this sample course guide) which are important to mastering each learner objective is given under the second heading. Each one represents understanding and/or skill to be acquired by the students.

Under the third heading, Sample Exercises and Solutions a variety of problems is presented and solutions are provided. In a few cases alternate methods of solution are shown. The performance objectives which each problem tests are given at the beginning of most exercises. "(P.O. 4, 16)" means that the student will develop or demonstrate mastery of Performance Objectives 4 and 16. The listing of sample exercises is just that—a sample. It is not intended to be complete. Not all objectives are represented and some are used in more than one sample exercise. While it is possible to incorporate some of these questions in tests and quizzes, the teacher will have to be selective, and supplementing will be necessary.

Suggested Student Reading and Film/Slide Presentations include journal articles, newspaper articles, books and chapters of books that students should be able to explore on their own or with a minimal amount of help from the teacher. This list can be used in a number of ways. If students are interested in statistics, they can be directed to some of these readings to expand their knowledge. Extra credit reports either in writing or as a class presentation can be prepared from these materials. Teachers can assign readings to individual students and time can be set aside in class for reports. In addition to readings, films and slide presentations which are appropriate to the topic are suggested. Information regarding ordering and company preview policy for these materials are in the audio-visual section of this guide.

Some of the information provided for the learner objectives is preceded by an asterisk (*). This indicates that the concept or performance objective is not required content, but that it is appropriate to the learner objective. The teacher should use his/her own judgment on how the information should be used. In some cases it may be introduced to the entire class as part of the course, it may not be dealt with at all, or it may be used to challenge the better students only.

I. LEARNER OBJECTIVE: The student distinguishes between descriptive and inferential statistics.

A. Important Concepts

1. Statistics is both a science and an art in which data are collected, organized, and analyzed to solve problems and to aid in decision making.

2. A Statistic is a single piece of data.

3. Descriptive Statistics is a branch of statistics used to describe or summarize data related to a specific population.

4. Inferential Statistics is a branch of statistics which uses information about a sample to make generalizations or predictions about the population when it is either impossible or impractical to gather all of the pertinent data.

Example: There is an election in October. Sheron Lee's name is on the ballot for the Board of Education. In September her supporters will use inferential statistics to analyze her standing relative to other candidates. After the election in October descriptive statistics will be used to analyze voter turnover and the election results.

5. A Population is the complete set of people, objects or events about which information is desired.

6. A Sample is a subset of the population.

Example: In September, when Sheron Lee's supporters try to analyze her standing in the political race, they will conduct a telephone survey. Using the list of registered voters her supporters will call a predetermined percentage of people on this list. The comprehensive list of registered voters makes up the population which is critical to this problem. The people from the list who are called make up a sample.

* 7. A Bias is a systematic error which affects all measurements the same way.

Example: An employee uses his car frequently for work and is reimbursed the cost of his mileage. If the odometer on his car is calibrated wrong, the employee's mileage reports will be biased.

B. Performance Objectives

1. The student can identify whether descriptive or inferential statistics has been used to reach a conclusion.

2. The student can determine whether descriptive or inferential statistics is appropriate for a given situation.

3. The student can identify the population in a problem.

4. The student can distinguish between population data and sample data.

5. The student can describe differences between descriptive and inferential statistics.

6. The student understands that risk is involved in using inferential statistics.

7. The student understands that the source of data is important.

* 8. The student understands why some statistical results are biased.

C. Sample Exercises

1. (P.O. 1,5) Find ten examples of statistics in the newspapers. Identify each as an example of descriptive or inferential statistics.

2. (P.O. 1,5) The word statistics is frequently used by radio and television newscasters. Listen to the news and classify each use as descriptive or inferential.

3. (P.O. 1,5) Classify each of the following according to use of descriptive or inferential statistics:

   a. A report from the Gallup Polls concerning an upcoming election.

   b. The Nielson ratings of television programs.

   c. The U.S. Census Report.

   d. Guinness Book of World Records.

   e. The Department of Labor's report on unemployment statistics.

4. (P.O. 1,4,7,8) Discuss the following quote from the National Smoker Study which appeared in a magazine advertisement:

   > "The overwhelming majority of menthol smokers reported that low tar Merit Menthol delivered taste equal to—or better than—leading high tar menthols. Cigarettes having up to twice the tar."

5. The following problem set from Statistics By Example is also an appropriate exercise for use with this objective:

"Tom Paine and Social Security" by William H. Kruskal and Richard S. Pieters in Exploring Data.

D. Suggested Student Reading and Film/Slide Presentations

1. Brown, B. W. Jr., "Statistics, Scientific Method and Smoking," in Statistics: A Guide to the Unknown.

2. Campbell, Stephen, Flaws and Fallacies in Statistical Thinking, Chapters 1, 2 and 3.

3. "Everything You've Always Wanted to Know About T.V. Ratings," A. C. Nielson Co., Northbrook, Illinois, 1974.

4. "Fact Finder for the Nation," film produced by the Census Bureau/U.S. Department of Commerce, available free through Film Services of Hawaii.

5. Hansen, Morris H., "How to Count Better: Using Statistics to Improve the Census," in Statistics: A Guide to the Unknown.

6. Huff, Darryl, How to Lie with Statistics, Chapters 1 and 7.

7. __________, How to Take a Chance, Chapter 8.

8. "The Science and Snares of Statistics," Time, Vol. 90 (Sept. 8, 1967): 29.

9. Taeuber, Conrad, "Information for the Nation from a Sample Survey," in Statistics: A Guide to the Unknown.

10. "We," film produced by Census Bureau/U.S. Department of Commerce, available free through Film Services of Hawaii.

11. "The Who, What, When, Where, Why of Polls," Honolulu Star-Bulletin, Friday, Nov. 3, 1978, pages A-1, A-4.

II. LEARNER OBJECTIVE: The student constructs frequency distributions and displays the data using various graphical methods.

A. Important Concepts

1. A Frequency Distribution is a convenient organization of data in which information is tallied in a table to indicate how often each data item or category occurs.

2. The Relative Frequency is a ratio which compares the frequency of a particular data item to the number of data items in the entire set of data.

3. A Class is a numerical grouping into which the data have been partitioned.

4. A Class Interval is a range of scores.

5. A Class Frequency represents the number of data items in a particular class.

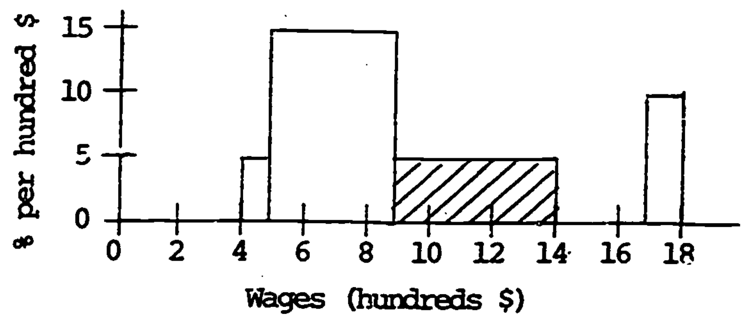
Example: Frequency Distribution of Students' Weight in Homeroom 8A

| Class | Interval | Midpoint | Tally | Frequency f | Relative Frequency r |
|---|---|---|---|---|---|
| 1 | 70–79 | 74.5 | /// | 3 | $\frac{3}{20} = 0.15$ |
| 2 | 80–89 | 84.5 | ~~////~~ // | 7 | $\frac{7}{20} = 0.35$ |
| 3 | 90–99 | 94.5 | ~~////~~ | 5 | $\frac{5}{20} = 0.20$ |
| 4 | 100–109 | 104.5 | /// | 3 | $\frac{3}{20} = 0.15$ |
| 5 | 110–119 | 114.5 | / | 1 | $\frac{1}{20} = 0.05$ |
| 6 | 120–129 | 124.5 | / | 1 | $\frac{1}{20} = 0.05$ |

$\Sigma r = 1.00$

6. A Histogram is a graph which represents continuous numerical data by area, not height. The area under one rectangle of a histogram (relative to the area of the entire histogram) corresponds to the relative frequency of the data it represents. A histogram does not need a vertical scale, although percent per unit is often used.

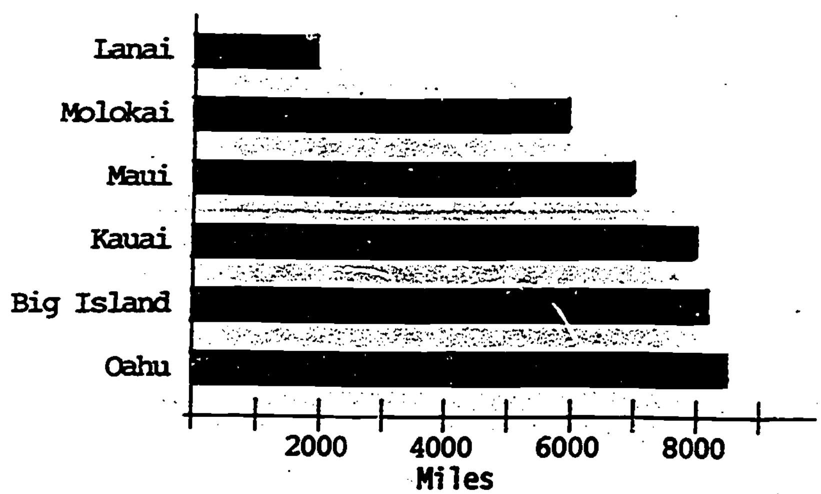
Example: Distribution of Monthly Wages




(The shaded area represents five one hundred dollar intervals; 25% of the monthly wages are between $900-$1,400.)

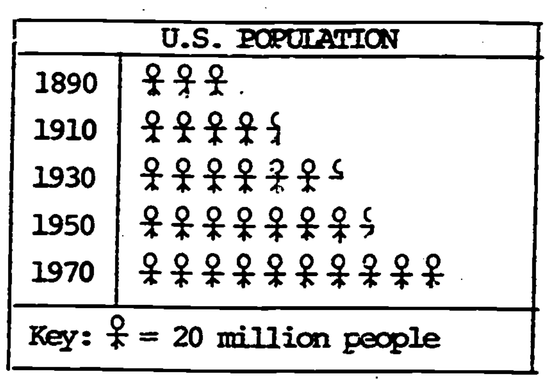
7. A Bar Graph uses vertical or horizontal bars of equal width to represent data. The class frequency is represented by the length of the bars.

Example: 1975 Yearly Mileage Per Car in Hawaii.



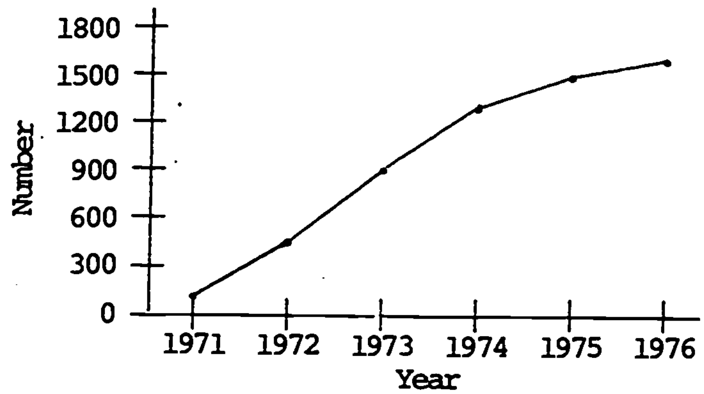

8. A Pictograph is a graph which uses pictures or symbols to represent the data. A key or legend indicating the number of items represented by each symbol must be given.

Example:



9. A Frequency Polygon, or Line Graph, is formed by imagining the bars of a histogram or bar graph and connecting the midpoints of these bars with straight lines. It is used to represent continuous data.

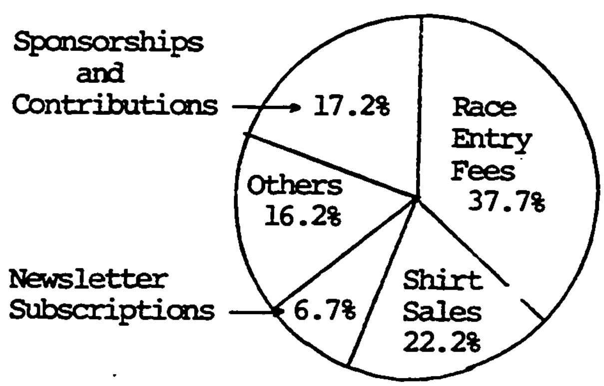
Example: Cable Television Subscribers in Kauai




Source: The State of Hawaii Data Rook, 1977

10. A Circle Graph, or Pie Chart, displays data and its component parts in a circular format. Each sector of the circle represents the proportional part of the data.

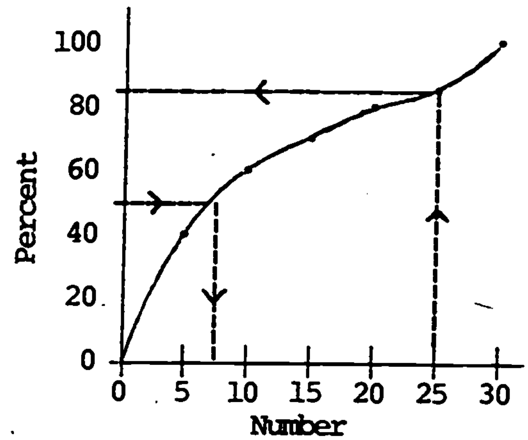
Example: Honolulu Marathon Association, 1978 Receipts




*11. An Ogive, or Cumulative Frequency Graph, is a graph in which one axis represents the values of a variable and the other axis represents cumulative frequency or percent. A point on the graph indicates the percent of observations which is either equal to or below that score.

Example: In the ogive on the next page the estimated 50th percentile is approximately score number 8. A score of 25 has an estimated percentile rank of 83.

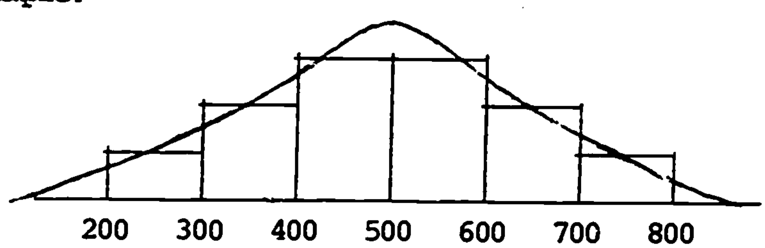
27




Scores on Math Exam

*12. A Normal Distribution is a symmetric distribution which has a bell-shaped curve.

Example:



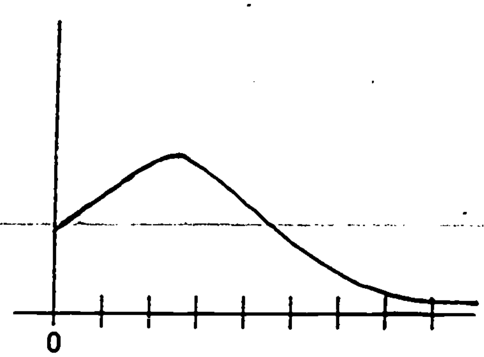

SAT MATH SCORES

*13. A Skewed Distribution is a distribution which is not symmetric.

*14. A Negatively Skewed Distribution has a relatively large number of large values.

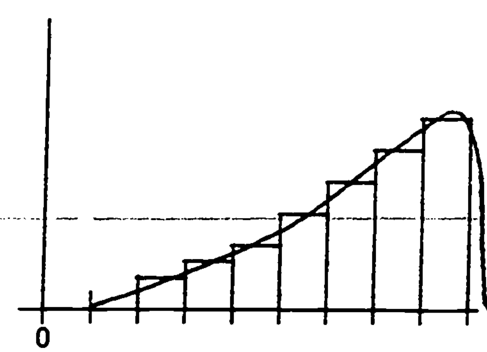
*15. A Positively Skewed Distribution has a relatively large number of small values.

Example:




Number of Children/Family in U.S.

(Positively Skewed Distribution)




Scores on an Easy Test

(Negatively Skewed Distribution)

B. Performance Objectives

1. The student can read data presented in tabular or graphical displays.

2. The student understands that to graphically display data, the data must first be organized and a frequency distribution compiled.

3. The student understands that the data in a frequency distribution should be organized in rank order.

4. The student can determine a reasonable number of intervals for a specific data set.

5. The student can determine a reasonable interval size and the class limits for each interval.

6. The student understands that each data item belongs in exactly one interval.

7. The student can group data into intervals.

8. The student understands why the number of intervals used to display data affects the appearance and possible interpretation of a graph.

9. The student understands the consequences of reducing the number of class intervals when displaying data.

10. The student can determine the relative frequency of an event or an occurrence.

11. The student understands that a histogram displays continuous numerical data.

12. The student understands that a histogram is preferable to a bar graph when the data is continuous.

13. The student understands that a key or legend must be given with a pictograph.

14. The student understands that histograms, bar graphs, and frequency polygons can be used to display data over time.

15. The student understands that a circle graph displays data from one point in time.

16. The student understands that a circle graph displays a distribution of parts.

17. The student can construct histograms, bar graphs, pictographs, circle graphs, and frequency polygons.

18. The student understands which type of graph is appropriate for displaying a specific set of data.

19. The student can compare two sets of similar data presented graphically.

20. The student understands that relative frequencies must be used to compare data sets with different numbers of observations.

21. The student can explain why certain graphs are misleading.

22. The student can describe how data has been distorted or misrepresented on certain graphs.

23. The student understands that the numerical scales of a graph should begin with zero.

24. The student understands that the numerical scales of a graph should be marked in equal intervals.

*25. The student can determine cumulative frequencies.

*26. The student can construct an ogive.

*27. The student can estimate the percentile rank of an observation from an ogive.

*28. The student can estimate which observation has a specific percentile rank from an ogive.

*29. The student can distinguish between normal and positively and negatively skewed distributions.

*30. The student can give examples of normal and positively and negatively skewed distributions.

C. Sample Exercises and Solutions

1. (P.O. 2-10,17) Audrey wanted to buy her boyfriend a camera for his birthday. Below are the prices, in dollars, she has been quoted by different stores for the same camera:

| | | | | |
|---|---|---|---|---|
| 319 | 285 | 395 | 405 | 350 |
| 355 | 370 | 347 | 296 | 399 |
| 324 | 412 | 315 | 350 | 350 |
| 340 | 360 | 333 | 389 | 425 |

a. Construct a frequency distribution.

b. Display the data on a histogram.

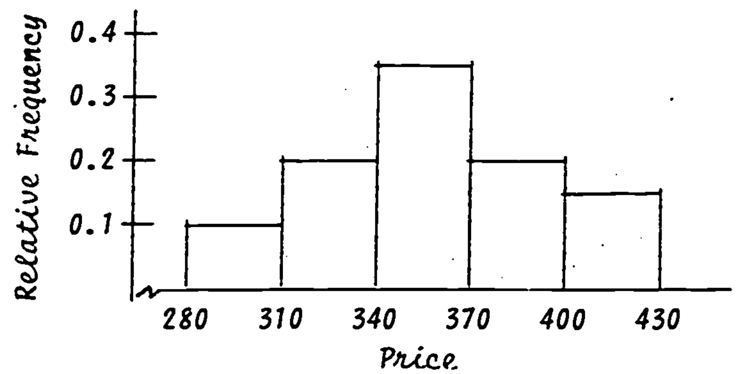
*Solution:*

a.

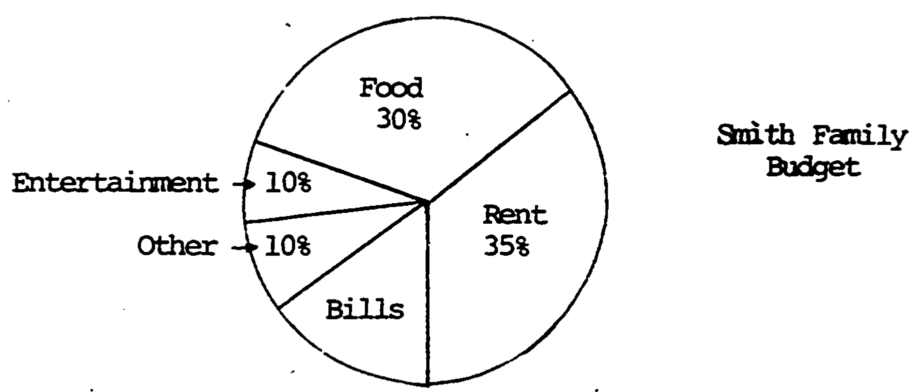
| *Price* | *Tally* | *Frequency* | *Relative Frequency* |
|---|---|---|---|
| *280-309* | *//* | *2* | *0.10* |
| *310-339* | *////* | *4* | *0.20* |
| *340-369* | *~~////~~ //* | *7* | *0.35* |
| *370-399* | *////* | *4* | *0.20* |
| *400-429* | *///* | *3* | *0.15* |

$\Sigma = 1.00$

b.




2. (P.O. 1,15,16,17)




a. What percent of the Smith's budget is spent on bills?

b. Is more money spent on food or rent?

c. If their monthly income is $700, how much do they spend on food?

*Solution:*

a. *15% [100% - (35% + 30% + 10% + 10%)]*

b. *Rent.*

c. *$210 (30% of $700 = $210)*

3. (P.O. 1,10,25) Complete the following summary table:

| Score | Frequency | Cumulative Frequency | Relative Frequency |
|---|---|---|---|
| 131-150 | 10 | 10 | ____ |
| 111-130 | ____ | 30 | ____ |
| 91-110 | 30 | ____ | ____ |
| 71-90 | ____ | ____ | 0.25 |
| 51-70 | 15 | ____ | ____ |
| 31-50 | ____ | 120 | ____ |

*Solution:*

| Score | Frequency | Cumulative Frequency | Relative Frequency |
|---|---|---|---|
| 131-150 | 10 | 10 | *0.08* |
| 111-130 | *20* | 30 | *0.17* |
| 91-110 | 30 | *60* | *0.25* |
| 71-90 | *30* | *90* | 0.25 |
| 51-70 | 15 | *105* | *0.13* |
| 31-50 | *15* | 120 | *0.13* |

4. (P.O. 21,22,23,24) Find examples of graphs in newspapers and magazines that are misleading. What causes these graphs to be misleading? Were they constructed this way deliberately?

5. Identify a relevant issue at your school or in your community and conduct a survey to gather opinions on this issue. Summarize the data collected in a frequency distribution and in an appropriate graph.

6. (P.O. 18) Which type of graph would be appropriate to display:

   a. The populations of Hawaii's major islands?

   b. The Hawaii County's harvest of macadamia nuts between 1950 and 1979?

   c. Sharen's Saturday morning weight from January 1979 to July 1979?

*Solution:*

*a.* *Bar Graph or Circle Graph or Pictograph.*

*b.* *Frequency Polygon (Line Graph) or Histogram.*

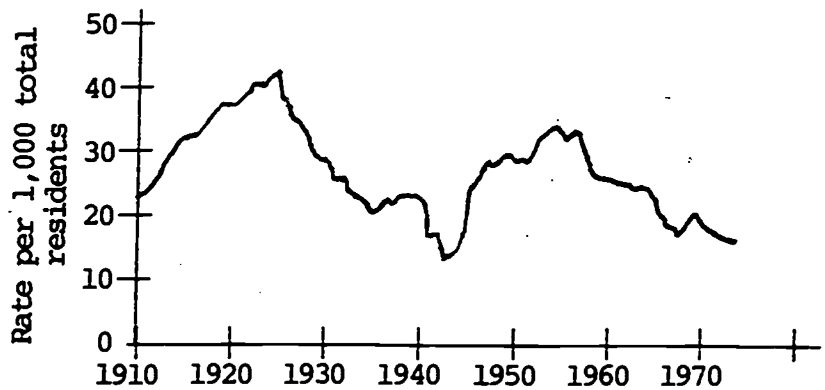
c. *Frequency Polygon.*

7. The following problem sets from Statistics By Example are also appropriate exercises for use with this objective:

   a. "Fractions on Closing Stock Market Prices" by Frederick Mosteller, Exploring Data.

   b. "Organizing and Reading Population Data" by Douglas H. Spicer, Exploring Data.

8. (P.O. 1,14) BIRTHS OCCURRING IN HAWAII




   a. How many births occurred during 1930?

   b. During which year were there approximately 38 births per 1,000 residents?

   c. Were there more births in 1940 or 1955?

   d. True or False: In 1970 approximately 22 births occurred in Hawaii.

   e. During which year between 1910 and 1970 did the fewest number of births occur?

   f. Why was the birthrate low at this time (see part e)?

   g. What is the current trend in terms of births occurring?

*Solution:*

*a.* *~28 per 1,000 residents*   *b.* *1920 and 1925*

c. *1955*   *d.* *False. ~22 per 1,000 residents*

e. *~1945* f. *WWII*

g. *Decreasing; more people are concerned with population control.*

9. Complete Chapter 9, Lessons 1, 2, and 5 in Harold Jacobs' book Mathematics: A Human Endeavor.

10. E T (A O) N (I R) (S H) D L (C U M) (F P Y) (W G B) V (K X J) (Q Z) This is a list of the 26 letters of the alphabet in decreasing order of use. Letters grouped by parentheses have about the same relative frequency in the English language.

    a. Analyze the Morse code in terms of this list. Do the symbols provide for an efficient language? Can you invent a better one?

    b. Analyze the game Scrabble in terms of this list? Does it make sense?

*11. (P.O. 29) Which of the following sets will have a normal distribution and which will be skewed? If skewed, indicate whether it is positive or negative.

    a. The age at which people retire.

    b. The annual income for families in the U.S.

    c. Temperature at 3 p.m. in Honolulu.

    d. IQ scores of people in the U.S.

    *Solution:*

    a. *Negatively skewed* b. *Positively skewed*

    c. *Normal* d. *Normal*

C. Suggested Student Readings and Film/Slide Presentations

1. Campbell, Steven, Flaws and Fallacies of Statistical Thinking, Chapter 5.

2. Huff, Darryl, How to Lie with Statistics, Chapters 5 and 6.

3. "Presentation of Data," slide/cassette presentation by Lansford Publishing Co.

4. "Statistics and Graphs I/II," video-cassette presentation by Great Plains National Instructional Television Library.

5. Tanur, Judith, editor, Statistics: A Guide to the Unknown.

III. LEARNER OBJECTIVE: The student computes percentile and percentile ranks.

A. Important Concepts

1. A Percentile is a measure of position within a distribution which gives the relative standing of an individual item as compared to the entire group of data. $P_i$ denotes that something is in the $i^{th}$ percentile.

2. A Percentile Rank of a term in a distribution indicates the percentage of scores within a group of data that fall below a specified score. It is found by adding the percentage of terms below the score, B, to one-half the percentage of terms equal to the score,

$$\text{Percentile Rank of } x = \frac{B + \frac{1}{2}E}{n} \cdot 100$$

Example: In one class scores on a math test were 27, 28, 31, 31, 42, 47, 48, 49, 49, 49, 49, and 53. Kimo's score of 49 has a percentile rank of 75. Another way of saying this is that Kimo scored in the 75th percentile.

$$(\text{Kimo's percentile rank} = \frac{7 + \frac{1}{2}(4)}{12} \cdot 100 = \frac{9}{12} \cdot 100 = 75)$$

3. A Raw Score is the value of numerical data given in the original units of measurement.

4. The Lower Quartile, $Q_1$, is the 25th percentile of a distribution. Twenty-five percent of the terms are below the quartile and 75 percent of the terms are above it.

5. The Upper Quartile, $Q_3$, is the 75th percentile of a distribution. Seventy-five percent of the terms are below $Q_3$ and 25 percent are above it.

Example: There are 20 terms in the following distribution of scores: 62, 68, 72, 73, 76, 76, 79, 83, 85, 85, 85, 86, 86, 88, 89, 91, 91, 93, 95, 100. 25 percent of 20 = 5, so $Q_1$ will lie between the 5th and 6th score: $Q_1 = 76$. 75 percent of 20 = 15, so $Q_3$ will lie between the 15th and 16th score: $89 < Q_3 < 91$.

B. Performance Objectives

1. The student understands why a raw score alone is often meaningless.

2. The student can determine a percentile rank for a score in a grouped data frequency distribution.

3. The student understands that a percentile rank is meaningless without reference to a specific group.

4. The student understands that a score with a percentile rank of N is the same thing as saying the score is in the Nth percentile.

5. The student understands that percentiles from two different groups cannot be compared without additional information.

6. The student understands that the 50th percentile is also called the median and the second quartile.

7. The student understands that the three quartile points divide the data into four equal parts.

8. The student understands that all histograms can be summarized using percentiles.

* 9. The student can determine the approximate raw score of a percentile in a grouped data frequency distribution.

10. The student understands that percentiles usually cannot be added, averaged or multiplied.

11. The student understands that percentiles are not evenly spaced along test score scales.

12. The student can determine percentages of a number.

C. Sample Exercises

1. Is it possible for a score to be in the 100th percentile? Give an example or explain why not.

*Solution:*
*It is impossible. The formula for finding the percentile for a score involves adding only one-half the percentage of terms equal to the score to the percentage of scores below the score. Therefore, the percentile of a score will always be less than 100: (100 - one-half the percentage of terms equal to the score).*

2. (P.O. 3,5) Maria and Stanford are in two different math classes which took the same final exam. Maria scored in the 78th percentile in her class and Stanford scored in the 83rd percentile in his. Who is the better student?

*Solution:*
*There is not enough information to answer the question. It is possible that Maria's class did exceptionally well and a raw score of 92 placed her in the 78th percentile. Stanford's class might not have done as well and a raw score of 71 might give him a percentile rank of 83.*

3. (P.O. 2,4,9) Twenty students at Hawaii High School took the College Board S.A.T. The mathematics scores for these

students are:

| | | | |
|---|---|---|---|
| Joe - 461 | Audrey - 386 | Hans - 555 | Allen - 479 |
| Bill - 286 | Eric - 579 | Mary - 320 | Kirk - 672 |
| Obe - 579 | Evans - 402 | Kate - 436 | Terry - 542 |
| Eseta - 461 | Maina - 315 | Todd - 584 | Dave - 337 |
| April - 395 | Tom - 463 | Thalia - 483 | Gail - 428 |

a. In this class what is the percentile rank of Terry?

b. What is the percentile rank of Audrey?

c. Which students scored closest to the 60th percentile?

*Solution:*

a. *There are 14 students who scored below Terry and there is one student (Terry herself) who equaled her score. Therefore,*

$$\text{Terry's percentile rank} = \frac{14 + \frac{1}{2}(1)}{20} \cdot 100 = 72.5$$

b. $$\text{Audrey's percentile rank} = \frac{4 + \frac{1}{2}(1)}{20} \cdot 100 = 22.5$$

c. *To be in the 60th percentile,*

$$\frac{B + \frac{1}{2}E}{20} \cdot 100 = 60 \text{ or } B + \tfrac{1}{2}E = 12$$

*When rank ordered, the scores that come closest to satisfying that requirement are Tom's and Allen's.*

*Tom's percentile rank = 57.5 and*
*Allen's percentile rank = 62.5*

4. (P.O. 2) In the junior class of 250 students only 55 students sold more than 25 Huli-Huli Chicken tickets. Malia and three other students each sold 25 tickets. Find Malia's percentile rank.

*Solution:*
*If 55 students sold more than 25 tickets and 4 students sold exactly 25 tickets, 250 - (55 + 4) = 191 students sold less than 25 tickets. Therefore,*

$$\text{Malia's percentile rank} = \frac{191 + \frac{1}{2}(4)}{250} \cdot 100 = 77.2$$

5. (P.O. 4,12) In a class of 32 students, Genoa had a percentile rank of 25. How many students had scores lower than Genoa?

*Solution:*
*Eight of the students had lower scores (25% of 32 = 8).*

6. (P.O. 2,4,9,12) This is a distribution of the heights of the ninth grade students at Kona High School:

| Height (in inches) | Frequency | Height (in inches) | Frequency |
|---|---|---|---|
| 54 | 1 | 60 | 9 |
| 55 | 1 | 61 | 7 |
| 56 | 3 | 62 | 5 |
| 57 | 4 | 63 | 2 |
| 58 | 7 | 64 | 2 |
| 59 | 8 | 65 | 1 |

a. What is the percentile rank of a height of 62 inches?

b. What percent of students are taller than 60 inches?

c. Roland is 58 inches tall. What is his percentile rank?

d. What percent of students are more than 60 inches tall, but less than 64 inches tall?

e. How tall is a student in the 40th percentile?

*Solution:*

a. *Forty of the 50 students have heights less than 62 inches. Therefore,*

$$\text{percentile rank} = \frac{40 + \frac{1}{2}(5)}{50} \cdot 100 = 85$$

b. *Seventeen of the students are taller than 60 inches (17/50 = 34%). (This question has nothing to do with percentile ranks!)*

c. *Roland's percentile rank* $= \frac{9 + \frac{1}{2}(7)}{50} \cdot 100 = 25$

d. *There are 14 students in this category. This is 14/50 or 28% of the students.*

e. *A student in the 40th percentile is taller than 40% of all students (40% of 50 = 20). Therefore, someone taller than 20 students, or someone 59 inches tall is in the 40th percentile.*

7. Have students find examples of percentiles and percentile ranks that affect them. (Examples include student rank upon graduation and percentile rank on College Board scores.)

D. <u>Suggested Student Readings</u>

1. Campbell, Stephen, <u>Flaws and Fallacies of Statistical Thinking</u>, Chapters 8 and 9.

2. Tanur, Judith, editor, <u>Statistics: A Guide to the Unknown</u>.

IV. LEARNER OBJECTIVE: The student calculates and compares the mean, median and the mode.

A. Important Concepts

1. The Mean is the arithmetic average of a set of numbers. It is found by dividing the sum of the numbers by the number of numbers added. $\bar{x}$ (read -bar) symbolizes the mean of the sample data and $\mu$, or mu, represents the population mean.

$$\text{Mean} = \frac{\Sigma x}{n}$$

$$\text{Mean for grouped data} = \frac{\Sigma xf}{\Sigma f}$$

2. The Median of a set of data is the middle value when the data is ordered. It is in the $(n+1)/2$ position, where n is the number of data elements. If an even number of data exist, the median is the arithmetic average of the middle two numbers.

3. In a set of numerical data, the Mode is the number which occurs most often. It is possible for a set to have one mode, no mode, or more than one mode.

Example: Given the set $A = \{1,1,2,3,4,4,5,8,9\}$

$$\text{Mean} = \mu_A = \frac{\Sigma x}{n} = \frac{1+1+2+3+4+4+5+8+9}{9} = \frac{37}{9} \approx 4.1$$

Median is the term in the $\frac{n+1}{2} = \frac{9+1}{2} = 5\text{th}$ position.

Therefore, Median = 4

This set is bimodal; the modes are 1 and 4.

Example: Given the set $B = \{1,2,3,4,1,4,3,8,9,4\}$

$$\text{Mean} = \mu_B = \frac{\Sigma x}{n} = \frac{39}{10} = 3.9$$

*Median is in the $\frac{n+1}{2} = \frac{10+1}{2} = 5\frac{1}{2}$ position.

Therefore, Median $= \frac{3+4}{2} = 3.5$

*Mode = 4

*To find the Median and Mode, the set had to be ordered: $B = \{1,1,2,3,3,4,4,4,8,9\}$.

4. Measures of Central Tendency are numerical methods that produce a single number to describe the location of a typical observation. Mean, median and mode are three measures of central tendency.

5. $\Sigma$ is the Greek symbol read as "sigma." It is used as a summation notation and appears in the formula for determining the mean.

$\Sigma x$ means "the sum of all x's."

6. The Weighted Mean is the arithmetic average of a set of numbers in which all values are not of equal importance. To compute a weighted mean multiply each number x by its assigned numerical weight w and then divide the sum of the products by the sum of the weights.

$$\text{Weighted mean} = \frac{\Sigma xw}{\Sigma w}$$

Example: Naomi has $1,000 earning 5% interest and $5,000 earning 8% interest. Therefore, her money is earning a mean interest of 7.5%.

$$\frac{\Sigma xw}{\Sigma w} = \frac{(0.05)1,000 + (0.08)5,000}{1,000 + 5,000} = 0.0075$$

B. Performance Objectives

1. The student understands that the word average can refer to the mean, median or mode of a set of numbers.
2. The student understands which statistic--mean, median or mode--is appropriate for a given situation.
3. The student understands that the mean is always a unique number.
4. The student understands that the mean can be calculated for any set of numerical data.
5. The student understands that in a histogram the mean acts as an equilibrium point, or "balance point."
6. The student understands that the mean is the most sensitive of the three averages because it is determined by every item in a set of data.
7. The student understands that extreme values distort the mean more than the median or the mode.
8. The student understands that when data are rank ordered the number of observations to the left of (or above) the median equals the number of observations to the right (or below).
9. The student understands that the median of a histogram separates the histogram into two sections with equal areas.
10. The student understands that the median is the second quartile of a distribution.
11. The student understands that the median can be used to describe events which cannot be quantified (e.g., something of average difficulty).

12. The student understands that to find the mode requires no calculations.

13. The student understands that the mode can be used to describe both quantitative and qualitative data.

14. The student understands the effects on the mean, median and mode when each piece of data is increased or decreased by a constant amount.

15. The student can find the overall mean of several sets of data given the mean and the number of observations in each set.

16. The student can read and use the summation notation.

*17. The student understands and can use subscripted variables.

18. The student can calculate the mean, median and mode for grouped data.

19. The student understands the relationship between the formula for finding the mean for ungrouped data and the formula for grouped data.

20. The student can find the mean of grouped data when class intervals are used in the frequency distribution.

21. The student understands that the mean value is an approximation when data are given in interval form.

22. The student can determine a weighted mean.

23. The student understands when it is appropriate to use a weighted mean.

*24. The student is aware of a special class of weighted averages called index numbers such as Gross National Product, Consumer Price Index and the NY Stock Exchange readings.

*25. The student can find the estimated median of grouped data when class intervals are used in the frequency distribution.

*26. The student understands the relationship of the mean, median and mode in a normal distribution.

*27. The student understands the relationship of the mean, median and mode in a skewed distribution.

28. The student can calculate the mean of a set of data by using an "assumed mean" (see Exercise C.1).

29. The student understands that the mean is the most useful measure of central tendency because it can be treated algebraically.

30. The student understands that people and companies try to influence others with information about averages that is often deceiving--it is unknown which average is meant.

C. <u>Sample Exercises and Solutions</u>

1. (P.O. 3,4,8,12,14,28) The number of cars in a parking lot each day this week was 48, 37, 49, 59, and 40.

   a. Compute the mean, median and modal number of cars.

   b. If there had been five additional cars parked in the lot each day, what would the mean and median be?

   *Solution:*

   *a. First arrange the data in rank order:*

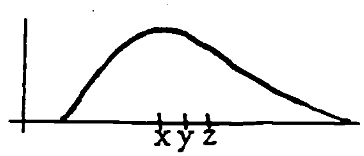
   *37, 40, 48, 49, 56*

   *Mean* $= \frac{\Sigma x}{n} = \frac{230}{5} = 46$; *Median* $= 48$; *there is no mode.*

   *a. (alternate solution for mean)*
   *Guess the value of the mean--45. The difference between each data point and the assumed mean of 45 is -8, -5, 3, 4, and 11. On the average, the assumed mean is wrong by* $\frac{-8+(-5)+3+4+11}{5} = \frac{5}{5} = 1$. *Therefore, the actual mean is* $45 + 1 = 46$.

   *b. Mean* $= 46 + 5 = 51$; *Median* $= 48 + 5 = 53$.

2. (P.O. 5,6,7,9,27) Use the words mean, median and mode to complete these statements.



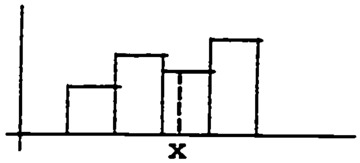

   a. x represents the _____ of the distribution.

   b. y represents the _____ of the distribution.

   c. z represents the _____ of the distribution.

   *Solution:*
   *a. mode* *b. median* *c. mean*
   *The mode, x, is the most frequently occurring value. The mean is most affected by the presence of extreme values in a distribution. The median separates the histogram into two sections of equal areas.*

4. (P.O. 5,8,9) The dotted line at x separates the histogram into two sections of equal area. What is x?




*Solution: The median.*

5. (P.O. 3,4,8,12,14,18,19,20,21,25) The test results in an introductory psychology class were as follows:

| Score | No. of Students |
|---|---|
| 70–79 | 40 |
| 60–69 | 75 |
| 50–59 | 50 |
| 40–49 | 20 |
| 30–39 | 15 |

a. The median and mode fall into which class intervals?

b. Calculate the approximate mean score.

*c. Calculate the approximate median score.

*Solution:*

a. *There are 200 students so the median is in the* $\frac{n+1}{2} = 100\frac{1}{2}$ *position. Therefore, it is in the interval 60-69. The mode is also in the interval 60-69.*

b. *The midpoints of the intervals are 74.5, 64.5, 54.5, 44.5, and 34.5. Using these values and the formula for grouped data:*

$$\mu = \frac{\Sigma xf}{\Sigma f} = \frac{74.5(40)+64.5(75)+54.5(50)+44.5(20)+34.5(15)}{200}$$

$$= \frac{11,950}{200} = 59.75$$

b. *(alternate solution)*
*Subtract 54.5 from each midpoint of the intervals: 20, 10, 0, -10, -20. Using these values and the formula for grouped data:*

$$\mu = 54.5 + \frac{\Sigma xf}{\Sigma f} = 54.5 + \frac{20(40)+10(75)+0(50)+-10(20)+-20(15)}{200}$$

$$= 54.5 + \frac{1,050}{200} = 54.5 + 5.25 = 59.75$$

*c. *The median is in the* $100\frac{1}{2}$ *position. Counting from the bottom there are 85 scores up to and including a score of*

43

*59. Another 15½ scores are needed in the next interval. Since there are 75 scores in that interval the median is the score $\frac{15\frac{1}{2}}{75}$ of the way between 60 and 69. Since $\frac{15\frac{1}{2}}{75} \times 10 \approx 2.1$ (ten is the length of the interval) the median must be 2.1 units beyond 59 (the last score counted). Therefore, the median is estimated to be 59 + 2.1 or 61.1.*

6. (P.O. 2,6,7,8) Which will be larger, the mean or the median age of guests at a local discothèque? Explain.

   *Solution:*
   *Probably the mean. Most discothèques have a minimum age requirement. Also, most guests tend to be in the same age categories. There are often a few "older" guests. The age of these "older" guests will affect the value of the mean much more than the value of the median.*

7. (P.O. 16,17) $y_1 = 7$, $y_2 = 13$, $y_3 = 14$, $y_4 = 0$, $y_5 = 6$, $\Sigma y =$ _____?

   *Solution:* $\Sigma y = 7 + 13 + 14 + 0 + 6 = 40$

8. (P.O. 3,4,18,19,22,23) Keola bought 20 shares of stock for $25, 50 shares of a second stock for $15 and 30 shares of a third stock for $20. What was his mean price for a share of stock?

   *Solution:*
   *Since each share of stock has a different value, or weight, the formula for weighted mean must be used.*

$$\mu = \frac{\Sigma xw}{\Sigma w} = \frac{20(25)+50(15)+30(20)}{60} = \frac{1,850}{60} \approx \$30.83$$

9. (P.O. 3,4,6) Leilani has worked for a total of 52 hours in the last three weeks. If she wants to average 16 hours a week, how many hours must she work this week?

   *Solution:*
   *This problem must be worked backwards. To average 16 hours for the four weeks, the total number of hours she must work is 16 × 4 = 64. Therefore, she must work 64 - 52 = 12 hours this week.*

10. Find examples of measures of central tendency used in newspapers, in magazines and on television. Have they been used properly? Are they deceiving?

11. The following problem sets from Statistics By Example are also appropriate for use with this objective:

    a. "Babies and Averages" by William H. Kruskal, Exploring Data.

b. "Characteristics of Members and Their Families" by Joseph I. Naus, Exploring Data.

c. "Collegiate Football Scores" by Frederick Mosteller, Exploring Data.

d. "The Cost of Eating" by William Kruskal, Exploring Data.

*e. "Prediction of Election Results from Early Returns" by Joseph Sedransk, Weighing Chances.

f. "Ratings of Typewriters" by Frederick Mosteller, Exploring Data.

12. Many of the student pages 513-539 from The Mathematics Resource Project: Statistics Information and Organization are also appropriate exercises for use with this objective.

13. Complete Chapter 9, Lesson 3 in Harold Jacobs' book Mathematics: A Human Endeavor.

D. Suggested Student Readings and Film/Slide Presentations

1. Angoff, William H., "Calibrating College Board Scores," in Statistics: A Guide to the Unknown.

2. Campbell, Stephen, Flaws and Fallacies in Statistical Thinking, Chapter 6.

3. "Distributions," sound/slide program by Prismatron Productions, Inc.

4. Huff, Darryl, How to Lie with Statistics, Chapters 2 and 3.

5. McCarthy, Philip J., "The Consumer Price Index," in Statistics: A Guide to the Unknown.

6. Moroney, M. J., Facts from Figures.

7. Phillips, John L., Jr., Statistical Thinking, A Structural Approach, Chapter 3.

8. "What is Average?" sound/slide program by Prismatron Productions, Inc.

45

V. LEARNER OBJECTIVE: The student "describes how scores are dispersed about the center by computing range, average deviation, variance, and standard deviation."

A. Important Concepts

1. The Range of a set of data is the difference between the largest and the smallest values in the distribution.

Example: For the second week in June, the high and low temperatures in Honolulu were as follows: 88/69, 87/68, 90/68, 90/70, 86/67, 89/70, and 85/68. Therefore, the range in temperature that week was 90° - 67° = 23°.

2. The Average Deviation of a data set is the sum of the absolute differences of the numbers from the mean divided by the number of observations. When describing sample data,

$$\text{average deviation} = \frac{\Sigma|x - \bar{x}|}{n}$$

and for population data,

$$\text{average deviation} = \frac{\Sigma|x - \mu|}{n}$$

3. The Variance of a set of data is the average of the squares of the differences of each observation from the mean. $\sigma^2$, or sigma squared, represents a population variance and the symbol $s^2$ is used for a sample variance. (Because $\Sigma(x - \bar{x})^2 \leq \Sigma(x - \mu)^2$ the divisor, n - 1, is used with sample data to avoid bias.)

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{n} \qquad \text{and} \qquad s^2 = \Sigma\,\frac{(x - \bar{x})^2}{n - 1}$$

More convenient computational equations for variance are:

$$\sigma^2 = \frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n^2} \qquad \text{and} \qquad s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n - 1}$$

Note: $\sigma^2$, the population variance, should not be confused with $\Sigma$, the small Greek symbol used as a summation notation.

4. The Standard Deviation describes the distance by which a typical observation differs from the mean. It equals the positive square root of the variance.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma(x - \mu)^2}{n}} \text{ and } s = \sqrt{s^2} = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

The corresponding computational equations for standard deviation are:

$$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n^2}} \qquad \text{and} \qquad s = \sqrt{\frac{n\Sigma x^2 - (\Sigma x)^2}{n(n - 1)}}$$

Rule of thumb: The ratio $\frac{\text{range}}{\text{standard deviation}}$ is rarely smaller than 2 or greater than 6.

5. Measures of Variation are methods of analyzing the dispersion or the differences among terms in a distribution. Range, variance, standard deviation, and average deviation are all measures of variation.

Example: $A = \{13,14,15,15,16,17\}$
$B = \{5,10,15,15,20,25\}$

In set A: $\mu = 15$, median = 15, mode = 15
In set B: $\mu = 15$, median = 15, mode = 15

Range of set A = 17 − 13 = 4
Range of set B = 25 − 5 = 20

$$\text{Avg. Deviation of A} = \frac{\Sigma|x - \mu|}{n} = \frac{2+1+0+0+1+2}{6} = \frac{6}{6} = 1$$

$$\text{Avg. Deviation of B} = \frac{\Sigma|x - \mu|}{n} = \frac{10+5+0+0+5+10}{6} = \frac{30}{6} = 5$$

$$\sigma_A^2 = \frac{\Sigma(x - \mu)^2}{n} = \frac{4+1+0+0+1+4}{6} = \frac{10}{6} = \frac{5}{3}$$

$$\sigma_B^2 = \frac{\Sigma(x - \mu)^2}{n} = \frac{100+25+0+0+25+100}{6} = \frac{250}{6} = \frac{125}{3}$$

$$\sigma_A = \sqrt{\sigma_A^2} = \sqrt{\frac{5}{3}} \approx 1.29 \quad ; \quad \sigma_B = \sqrt{\sigma_B^2} = \sqrt{\frac{125}{3}} \approx 6.45$$

In every case, the measures of variation for set B are larger than those for set A. The data in set B is more dispersed or spread out.

*6. Coding is a method of simplifying computations for the mean, variance and standard deviation by first adding or subtracting a constant amount to each term in a distribution.

Example: Five light bulbs with an "average life of 1,000 hours" were tested by the quality control department. They burned for 960, 1,070, 995, 1,040 and 980 hours each. To find the variance and standard deviation, the data is arranged in increasing order and coded by subtracting 1,000 from each number. The variance and standard deviation remain the same for coded data as for original data.

| | $x$ | $x^2$ |
|---|---|---|
| | −40 | 1,600 |
| | −20 | 400 |
| | −5 | 25 |
| | +40 | 1,600 |
| | +70 | 4,900 |
| sum: | 45 | 8,525 |

Therefore, $\sigma^2 = \frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n^2}$

$= \frac{8,525}{5} - \frac{45^2}{25}$

$= 1,705 - 81 = 1,624$

$\sigma = \sqrt{1,624} \approx 40.3$



47

(The mean is obtained by adding 1,000 to the mean of the new distribution.

Therefore, $\mu = \frac{\Sigma x}{5} + 1,000 = \frac{45}{5} + 1,000 = 1,009.$)

B. Performance Objectives

1. The student understands that the range depends solely on the two extreme values in a distribution.

2. The student understands that the range is affected by extreme values.

3. The student understands that the range does not indicate how the other terms vary.

4. The student can determine the square root of a number by hand, with a calculator or by using a table of square roots.

5. The student understands that the sum of the differences of each observation from the mean is zero. $[\Sigma(x - \bar{x}) = 0$ and $\Sigma(x - \mu) = 0]$

6. The student understands that in most cases the measures of variation use a measure of central tendency as a reference.

7. The student understands that the variance measures the mean of the squared deviations of each observation from the mean.

8. The student can determine the range, average deviation, variance and standard deviation for grouped data.

9. The student can determine the range, average deviation, variance and standard deviation for grouped data when class intervals are used in the frequency distribution.

10. The student understands that the variance, the standard deviation, and the average deviation are measures of variability that are affected by every observation.

11. The student understands that the larger the value of $\sigma$, the standard deviation, the more dispersed the scores are about the mean.

12. The student understands that $\sigma$ must have a value between the largest deviation from the mean and the smallest deviation from the mean.

*13. The student understands that for many sets of data, roughly 68% of the observations are less than one standard deviation away from the mean and that roughly 95% are less than two standard deviations away.

14. The student can explain the concept of variability and the purpose of the measures of variability.

15. The student can describe the differences in two sets of data with identical measures of central tendency, but with different standard deviations.

16. The student understands that if each data item is increased or decreased by a constant amount, the standard deviation and the variance are unchanged.

*17. The student understands that coding can simplify computations if used correctly.

18. The student understands that the average deviation is less popular than the standard deviation because absolute value is difficult to work with algebraically.

C. Sample Exercises and Solutions

1. (P.O. 10,11,12) What does a distribution look like if the standard deviation is zero?

*Solution:*
*Every element of the set is equal to the mean.*

$$x_1 = x_2 = x_3 = \ldots = x_n = \mu$$

2. (P.O. 4,16) Given A = {5,6,6,8,9,9,13}

a. Find the mean and standard deviation.
b. Create set B by adding 12 to each element of set A.
c. Find the mean and standard deviation of set B.
d. Create set C by subtracting 4 from each element of set A.
e. Find the mean and standard deviation of set C.
f. If set D is formed by adding 21 to each element of set A, what will the mean and standard deviation of set D be?

*Solution:*

| | Set A | | Set B | | Set C | |
|---|---|---|---|---|---|---|
| | $x$ | $x^2$ | $y$ | $y^2$ | $z$ | $z^2$ |
| | 5 | 25 | 17 | 289 | 1 | 1 |
| | 6 | 36 | 18 | 324 | 2 | 4 |
| | 6 | 36 | 18 | 324 | 2 | 4 |
| | 8 | 64 | 20 | 400 | 4 | 16 |
| | 9 | 81 | 21 | 441 | 5 | 25 |
| | 9 | 81 | 21 | 441 | 5 | 25 |
| | 13 | 169 | 25 | 625 | 9 | 81 |
| *sum:* | 56 | 492 | 140 | 2,844 | 28 | 156 |

a. $\mu_A = \frac{\Sigma x}{n} = \frac{56}{7} = 8$

$$\sigma_A = \sqrt{\frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n^2}}$$

$$= \sqrt{\frac{492}{7} - \frac{56^2}{49}}$$

$$\approx \sqrt{6.29} \approx 2.5$$

c. $\mu_B = \frac{\Sigma y}{n} = \frac{140}{7} = 20$ ; $\sigma_B = \sqrt{\frac{2,844}{7} - \frac{140^2}{49}} \approx \sqrt{6.29} \approx 2.5$

e. $\mu_C = \frac{\Sigma z}{n} = \frac{28}{7} = 4$ ; $\sigma_C = \sqrt{\frac{156}{7} - \frac{28^2}{49}} \approx \sqrt{6.29} \approx 2.5$

f. $\mu_D = \mu_A + 21 = 8 + 21 = 29$ ; $\sigma_D = \sigma_A \approx 2.5$

3. (P.O. 4,16) Given A = {5,6,6,8,9,9,13} (the same set as in 2 above).

   a. Create set E by multiplying each element of set A by 2.
   b. Guess the values of the mean and standard deviation of set E, then compute the actual values.
   c. Explain your results.

   *Solution:*

   a. E = {10,12,12,16,18,18,26}

   b. $\Sigma x = 112$ ; $\Sigma x^2 = 2{,}292$

   $\mu_E = \frac{\Sigma x}{n} = \frac{112}{7} = 16$ ; $\sigma_E = \sqrt{\frac{2{,}292}{7} - \frac{112^2}{49}} \approx \sqrt{71.4} \approx 8.45$

   c. *When each element of a set is multiplied by a constant, the mean of the new set equals the mean of the old set times the same constant. The standard deviations of the two sets will not be equal because each data item has been increased by a different amount.*

4. (P.O. 4,16) In 1975, Gov. Brown of California suggested that all state employees be given a flat $70 a month raise. How would this affect the average monthly salary of the employees? How would it affect the standard deviation?

   *Solution:*
   *The average monthly salary (mean) would be increased by $70. The standard deviation would remain unchanged.*

5. (P.O. 4,12,16) In Hawaii in 1979, the state offered some of its bargaining units a 7% raise, across the board. How would this affect the mean and standard deviation of their salaries?

   *Solution:*
   *The mean salary would be 7% higher. The standard deviation will be larger. The top-salaried employees will get larger raises than anyone else and this will make the difference between their salaries and the mean salary larger than it was before the raise.*

6. (P.O. 12) Can the standard deviation ever be negative? Explain.

   *Solution:*
   *No. By definition the standard deviation is always non-negative.*

7. (P.O. 12,15) Each of the following lists has an average of 30. Guess whether the standard deviation for each is closer to 1, 2 or 10.

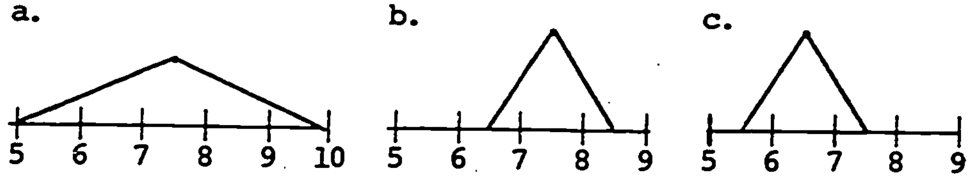
a. 29, 31, 29, 31, 29, 31, 29, 31, 29, 31
b. 28, 32, 28, 32, 28, 32, 28, 32, 28, 32
c. 28, 31, 29, 32, 27, 32, 26, 31, 33, 31
d. 34, 29, 26, 29, 31, 33, 30, 30, 29, 29
e. 40, 16, 11, 30, 28, 30, 34, 36, 42, 33

*Solution:*
*a. 1 b. 2 c. 2 d. 2 e. 10*

8. (P.O. 6,11,13,15) Which of the following descriptions matches the sketches of the histograms below?

(i) mean = 6.5; SD = 1
(ii) mean = 6.5; SD = 0.5
(iii) mean = 7.5; SD = 1
(iv) mean = 7.5; SD = 0.5
(v) mean = 7.5; SD = 2
(vi) mean = 7.5; SD = 1.5

a. b. c.




*Solution:*
*Roughly 95% of the observations should be less than two standard deviations from the mean. Therefore:*

*a. (iii) $\mu = 7.5$ ; $\sigma = 1$ b. (iv) $\mu = 7.5$ ; $\sigma = 0.5$*

*c. (ii) $\mu = 6.5$ ; $\sigma = 0.5$*

9. (P.O. 1,4) LiAnn loves to talk on the phone. Tuesday she spoke for 22, 36, 18, 27, 41 and 15 minutes during six different calls. Find the range, mean, variance, standard deviation, and average deviation for the length of a telephone call.

*Solution:*

| | $x$ | $x^2$ |
|---|---|---|
| | 15 | 225 |
| | 18 | 324 |
| | 22 | 484 |
| | 27 | 729 |
| | 36 | 1,296 |
| | 41 | 1,681 |
| *sum:* | 159 | 4,739 |

*Range* $= 41 - 15 = 26$

*Mean* $= \frac{159}{6} = 26.5$

*Variance* $= \frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n^2} = \frac{4{,}739}{6} - \frac{159^2}{36} \approx 87.6$

*Standard deviation* $= \sqrt{\textit{variance}} \approx 9.4$

*Average deviation* $= \frac{\Sigma|x - \mu|}{n}$

$$= \frac{11.5+8.5+4.5+0.5+9.5+14.5}{6} = \frac{49}{6} \approx 8.2$$

10. (P.O. 1,10) Which of the following measures of variability, if any, is not dependent on the exact value of each score?

a. range
b. average deviation
c. variance
d. standard deviation

*Solution:*
a. *The range is only dependent on the two extreme values.*

11. (P.O. 7,9) Find the variance for the following distribution:

| Interval | Frequency |
|---|---|
| 20-24 | 6 |
| 25-29 | 8 |
| 30-34 | 5 |
| 35-39 | 10 |

*Solution:*
*The variance must be estimated using the interval midpoints as x values.*

| *Interval* | $x$ | $f$ | $xf$ | $x^2$ | $x^2f$ |
|---|---|---|---|---|---|
| 20-24 | 22 | 7 | 154 | 484 | 3,388 |
| 25-29 | 27 | 8 | 216 | 729 | 5,832 |
| 30-34 | 32 | 5 | 160 | 1,024 | 5,120 |
| 34-39 | 37 | 10 | 370 | 1,369 | 13,690 |
| *sums:* | | 30 | 900 | | 28,030 |

*For grouped data,*

$$\sigma^2 = \frac{\Sigma x^2 f}{n} - \frac{(\Sigma x f)^2}{n^2} = \frac{28,030}{30} - \frac{(900)^2}{30^2} = 934.\overline{3} - 900 = 34.\overline{3}$$

12. (P.O. 13) A cylindrical part for the Saturn rocket is to have a diameter of 45 mm. The part has a standard deviation of 0.5 mm in the production process. If NASA will only accept parts with diameters in the range of 44.5 to 45.5 mm, what percent of parts will be rejected?

*Solution:*
*NASA will accept parts in the range 44.5 to 45.5 mm. Another way to state this is 45 ± 0.5 mm. The acceptable parts must be within one standard deviation of 45 mm. This means roughly 68% of the parts will be accepted and 32% will be rejected.*

*13. The following problem set from Statistics by Example is an appropriate exercise for use with this objective for students who have extended their work to include the area under the normal distribution:

"How Much Does a 40-lb Box of Bananas Weigh?" by Ralph B. D'Agostino in Detecting Patterns.

14. Many of the student pages 541--569 from The Mathematics Resource Project: Statistics Information and Organization are also appropriate exercises for use with this objective.

15. Complete Chapter 9, Lesson 4 in Harold Jacobs' book Mathematics: A Human Endeavor.

D. Suggested Student Reading and Film/Slide Presentation

1. Campbell, Stephen, Flaws and Fallacies in Statistical Thinking, Chapter 7.

2. "Descriptive Statistics," a sound/slide presentation available from Lansford Publishing Company.

3. Huff, Darryl, How to Lie with Statistics, Chapter 4.

4. Phillips, John L. Jr., Statistical Thinking, A Structural Approach, Chapter 4.

5. Tanur, Judith, editor, Statistics: A Guide to the Unknown.

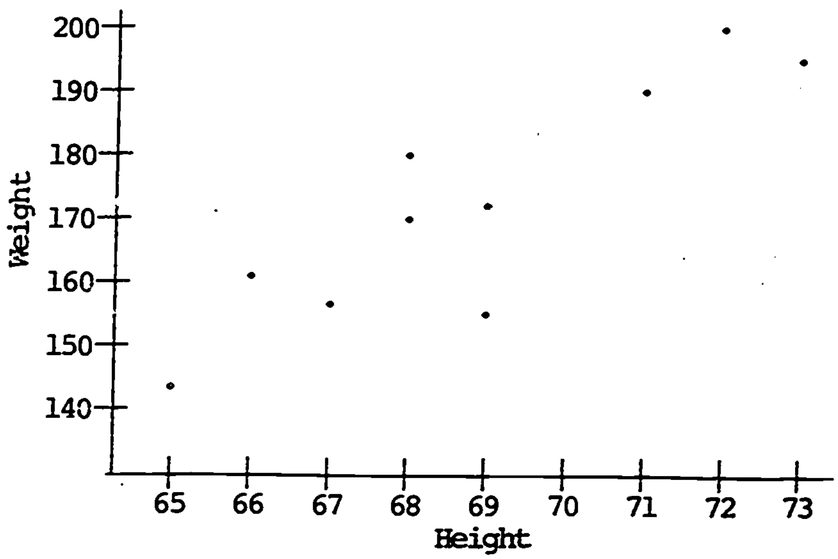
VI. LEARNER OBJECTIVE: The student calculates linear correlation coefficients and draws a scattergram.

A. Important Concepts

1. A Scatter Diagram is a graphical technique used to analyze the relationship between two variables. Each of the two axes of the graph represents one variable. Points on the graph represent paired observations of both variables.

Example: The following data was recorded for 10 adult males:

| Height (in inches) | 68 | 66 | 71 | 69 | 69 | 68 | 65 | 72 | 67 | 73 |
|---|---|---|---|---|---|---|---|---|---|---|
| Weight (in pounds) | 170 | 161 | 190 | 172 | 155 | 180 | 143 | 200 | 158 | 195 |




2. A Linear Correlation is a relationship between two variables which tells whether an increase in the value of one variable will cause an average change in the value of the other variable. If two variables have a linear correlation, the points on the corresponding scatter diagram will form an approximate straight line.

Example: The variables, height and weight, have a linear correlation as seen in the above scatter diagram.

3. The Coefficient of Linear Correlation, r, is a statistic that measures the strength of the linear relation between two variables. The computational formula for r is:

$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2} \cdot \sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

It will always be true that $-1 \le r \le +1$. If $|r|$ is close to 1, the correlation is said to be strong. If $|r|$ is close to 0, there appears to be no correlation.

Example: Using the above example of heights and weights, $\Sigma x = 688$, $\Sigma x^2 = 47,394$, $\Sigma y = 1,724$, $\Sigma y^2 = 300,368$ and $\Sigma xy = 118,995$. Therefore,

$$r = \frac{10(118,995) - 688(1,724)}{\sqrt{10(47,394) - 688^2} \cdot \sqrt{10(300,368) - 1,724^2}}$$

$$= \frac{3,838}{\sqrt{596} \cdot \sqrt{31,504}} \approx \frac{3,838}{(24.4) \quad (177.5)} \approx 0.89$$

Since r is very close to 1, this indicates there a high correlation between height and weight in adult men.

4. A Positive Correlation indicates that the two variables are directly related. That is, as one variable increases so does the other.

   Example: There is a positive correlation between height and weight ($r = +0.89$). This indicates that as men get taller, their weight also rises.

5. A Negative Correlation indicates that there is an inverse relation between two variables. That is, as x increases, y decreases.

   Example: There is a negative correlation between the age of a car and its selling price (excluding antiques). The older a car gets, the less it costs.

*6. The Level of Significance indicates the percentage of times for which a correlation is statistically significant. A chart or table in most statistics texts helps to interpre+ the value of the correlation coefficient correctly.

   Example: Using a table on the Critical Value of r indicates that there is a definite positive correlation between height and weight over 99% of the time. (For $n = 10$, $r_{0.005} = 0.765$. This means that if $|r| < 0.765$, there is no statistical significance to the correlation 99% of the time.)

B. Performance Objectives

1. The student understands that in a scatter diagram the independent variable is graphed on the horizontal axis.

2. The student understands that in a scatter diagram the scales do not necessarily begin with zero.

3. The student can plot points on a two-dimensional graph.

4. The student can interpret points plotted on a two-dimensional graph.

5. The student understands that the correlation coefficient will always be between -1 and +1.

6. The student understands that there is a perfect correlation, that is all points lie on a straight line, if $|r| = 1$.

7. The student understands that there is a high linear correlation if $|r|$ is close to 1.

8. The student understands that there is a low linear correlation if $|r|$ is close to zero.

9. The student understands that there is no linear correlation if $r = 0$.

10. The student understands that if $r \approx 0$, there is essentially no linear relationship, but another relationship may exist.

11. The student can predict whether two variables will have a positive, negative, or zero correlation.

12. The student can explain the purpose and fundamental characteristics of the correlation coefficient.

13. The student understands that $r = 2x$ does not indicate twice as much linearity as $r = x$.

14. The student understands that correlation measures association, but it does not measure causation.

15. The student understands that because it determines whether two variables are related, but does not specify how, the correlation coefficient cannot be used to solve prediction problems.

16. The student understands that the correlation coefficient is one of the most misused statistical measures.

17. The student understands that even though two things appear to be correlated, there may be a third variable which is highly correlated to both.

18. The student understands that the correlation coefficient indicates the direction and degree of the relation between two variables.

*19. The student understands that the correlation coefficient is unaffected by coding.

*20. The student understands that the correlation coefficient is unaffected if the two variables are coded in different ways.

*21. The student can determine whether a significant statistical correlation exists between two variables.

## C. Sample Exercises and Solutions

1. (P.O. 9) Indicate whether each of the following will have a positive, negative, or zero correlation:

   a. A parent's IQ and a child's IQ.
   b. The price of gas and the number of economy cars sold.
   c. The number of hours studied and grade on exam.
   d. The number of students in New Hampshire and the number of pets in Vermont.
   e. The interest rate on mortgages and the number of new home loans issued.

   *Solution:*
   *a. positive*    *b. positive*    *c. positive*
   *d. zero*    *e. negative*

2. (P.O. 5) Which set of paired observations shows a stronger linear relation: one in which $r = -0.8$ or $r = +0.7$?

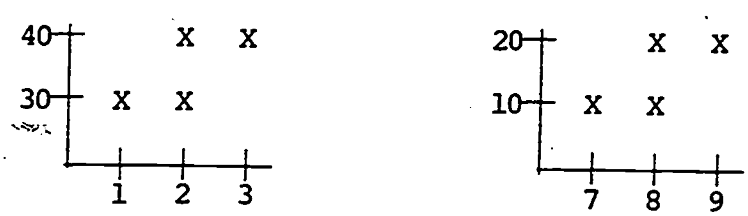
   *Solution:*
   *$|-0.8| > |+0.7|$, therefore the first one has a higher correlation.*

3. Give examples from everyday experience that demonstrate a correlation between variables.

4. (P.O. 4,9) If daughters were always 10% shorter than their fathers, what would the correlation of their heights be?

   *Solution:*
   *A perfect positive correlation would exist. Therefore, $r = +1$.*

5. (P.O. 10,17,18) How are the correlation coefficients related in the scatter diagrams below?



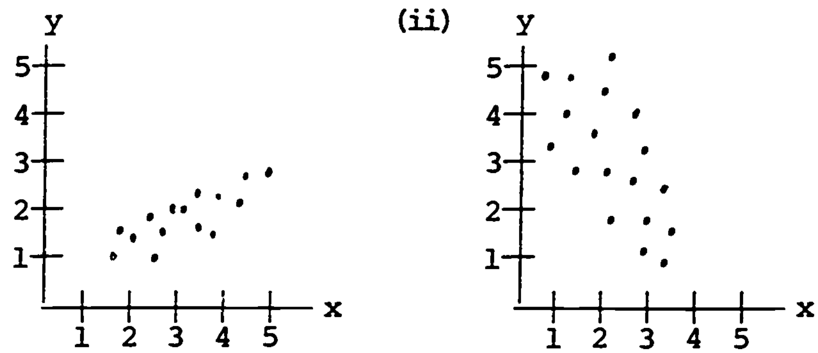

   *Solution:*
   *The correlation coefficients are the same. The graphs just have a change in scale.*

6. (P.O. 12,14,15) There is a positive correlation between blood pressure and income. How can this be explained? Is there a causal connection?

*Solution:*
*A positive correlation does not imply cause and effect. There may be a third variable, such as age, which is highly correlated to both.*

7. (P.O. 4,9) Study each of the scatter diagrams below:

(i) (ii)




Answer the following questions for each one:

a. Is the average of x closest to 2, 3, or 4?
b. Is the average of y closest to 2, 3, or 4?
c. Is the standard deviation of x closest to 0.5, 0.75, or 1?
d. Is the standard deviation of y closest to 0.3, 1, or 1.5?
e. Is the correlation positive, negative or zero?

*Solution:*
*For (i):*
- *a.* *3* *b.* *2*
- *c.* *0.75 (mean ± 2 S.D. should include approximately 95% of the observations)*
- *d.* *0.3* *e.* *positive*

*For (ii):*
- *a.* *2* *b.* *3*
- *c.* *0.5* *d.* *1*
- *e.* *negative*

8. (P.O. 5,9,16) The correlation between the heights of fathers and mature sons is:

(a) exactly −1; (b) close to −1; (c) close to 0;
(d) close to +1; (e) exactly +1.

*Solution:*
*(d) There is a definite positive correlation, but it is not perfect. Therefore, it is close to +1.*

9. (P.O. 5,18) Find the correlation coefficient for ten months of a company's monthly advertising expenditure and its total monthly sales. If x = advertising expenditures and y = total monthly sales, $\Sigma x = 426$, $\Sigma y = 735$, $\Sigma x^2 = 18{,}416$, $\Sigma y^2 = 54{,}803$ and $\Sigma xy = 31{,}763$.

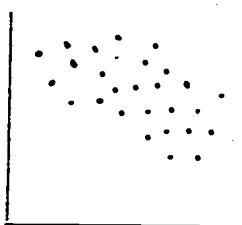
*Solution:*

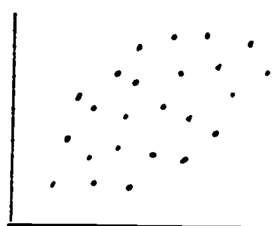
$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2} \cdot \sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

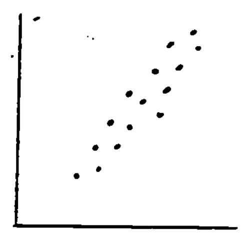
$$= \frac{10(31,763) - 426(735)}{\sqrt{10(18,416) - 426^2} \cdot \sqrt{10(54,803) - 735^2}}$$

$$= \frac{4,520}{\sqrt{2,684} \cdot \sqrt{7,805}} \approx \frac{4,520}{52 \cdot 88} \approx 0.99$$

10. (P.O. 5,6,9,16) The correlation coefficients for the three scatter diagrams below are 0.8, 0.2, and −0.6, not necessarily in that order. Match the diagrams with the correct value.

a. 

b. 

c. 

*Solution:*
*a. −0.6* *b. 0.2* *c. 0.8*

11. (P.O. 5,18,19) Twelve students recently took a math test and a science test. Their grades are shown below. Find the correlation coefficient for these grades.

| (x) Math | 83 | 71 | 48 | 95 | 65 | 88 | 63 | 72 | 100 | 77 | 35 | 69 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Science (y) | 69 | 74 | 72 | 91 | 72 | 94 | 67 | 73 | 98 | 86 | 43 | 77 |

*Solution:*
$\Sigma x = 866$, $\Sigma x^2 = 66,296$, $\Sigma y = 896$, $\Sigma y^2 = 69,838$ *and* $\Sigma xy = 67,791$.
*Therefore,*

$$r = \frac{12(67,791) - 866(896)}{\sqrt{12(66,296) - 866^2} \cdot \sqrt{12(69,838) - 896^2}}$$

$$= \frac{37,556}{\sqrt{45,596} \cdot \sqrt{35,240}} \approx \frac{37,556}{214 \cdot 188} \approx 0.93$$

*Alternate Method of Solution:*
*First code both sets of scores by subtracting 75 from each grade:*

| x | 8 | −4 | −27 | 20 | −10 | 13 | −12 | −3 | 25 | 2 | −40 | −6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | −6 | −1 | −23 | 16 | −3 | 19 | −8 | −2 | 23 | 11 | −32 | 2 |

$\Sigma x = -34$, $\Sigma x^2 = 3,896$, $\Sigma y = -4$, $\Sigma y^2 = 2,938$ *and* $\Sigma xy = 3,141$.

$$r = \frac{12(3,141) - (-34)(-4)}{\sqrt{12(3,896) - (-34)^2} \cdot \sqrt{12(2,938) - (-4)^2}}$$

$$= \frac{37,556}{\sqrt{45,596} \cdot \sqrt{35,240}} \approx 0.93$$

12. (P.O. 7,18,21) What is the difference between a situation where r = −0.8 and one where r = +0.8? Which will result in better predictions?

    *Solution:*
    *Both situations are equally highly correlated. However, one has a positive correlation and the other is negative. Before answering which will result in a better prediction, more information (including level of significance) is needed.*

13. The following problem set from Statistics By Example is also an appropriate exercise for use with this objective:

    "Examples of Graphical Methods" by Yvonne M. M. Bishop in Exploring Data.

14. Some of the exercises in student pages 483-496 in the Mathematics Resource Project: Statistics and Information Organization are also appropriate to this objective.

D. Suggested Student Readings

1. Andersen, Leonall C., "Statistics for Public Financial Policy" in Statistics: A Guide to the Unknown.
2. Brown, B. W. Jr., "Statistics, Scientific Method, and Smoking" in Statistics: A Guide to the Unknown.
3. Campbell, Stephen K., Flaws and Fallacies in Statistical Thinking, Cahpters 10 and 13.
4. "Gun Use Drops in Massachusetts," Honolulu Star-Bulletin, June 30, 1979, page F-3.
5. Huff, Darryl, How to Lie with Statistics, Chapter 8.
6. Link, Richard F., "Election Night on Television" in Statistics: A Guide to the Unknown.
7. "Reef Runway Has Most Tire Blowouts," Honolulu Star-Bulletin, June 21, 1979, pages A-1 and A-4.
8. Whitney, C. A., "Statistics, the Sun, and the Stars" in Statistics: A Guide to the Unknown.

VII. LEARNER OBJECTIVE: The student predicts one variable's value, knowing another, by using the least squares regression equation and line.

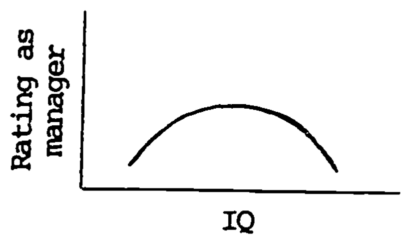
A. Important Concepts

1. Regression Methods are methods designed to find the best fit of an equation to a set of numbers. These methods make it possible to predict one variable's value from the observation of the other.

2. Linear Regression is the method used to fit a straight line to a set of numbers.

*3. Curvilinear Regression fits a curve to a set of numbers.

Example:

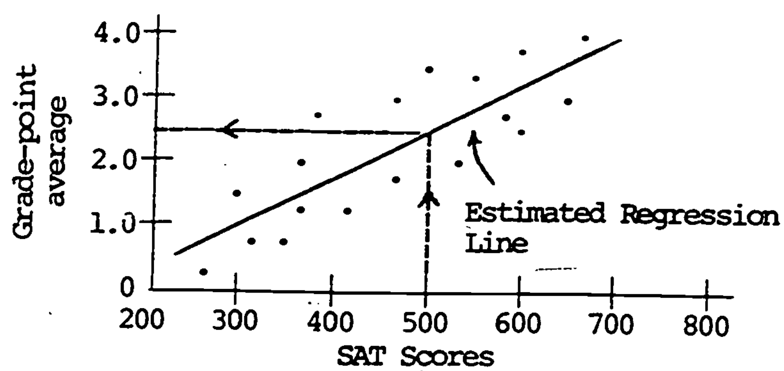



4. An Estimated Regression Line is a trend line which best represents the two variables on a scatter diagram. The points on the regression line are used to estimate the value of the dependent variable given the value of the independent variable.

5. The Dependent Variable is the variable for which the value is to be predicted.

6. The Independent Variable is the variable for which the value is known.

Example:




The independent variable is SAT scores and the grade-point average is the dependent variable. A student who scores 500 on the SAT is predicted to have a 2.5 GPA.

7. The Least Squares Method is a regression method used to determine the estimated regression line. It identifies the linear

equation so that the sum of the squared distances between the line and the data points is as small as possible. The equation of the estimated regression line is:

$$y_{predicted} = \bar{y} + b(x - \bar{x}) \quad \text{where}$$

$\bar{x}$ and $\bar{y}$ are the sample means of x and y and

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2} \quad \text{where}$$

n = the number of paired observations.

Example: Using the data from the example given in Learner Objective VI on heights and weights of adult males:

$$b = \frac{10(118{,}995) - (688)(1{,}724)}{10(47{,}394) - 688^2} = \frac{3{,}838}{596} \approx 6.4$$

$$\bar{x} = \frac{688}{10} = 68.8 \qquad \bar{y} = \frac{1{,}724}{10} = 172.4$$

Therefore, the estimated regression line is

$$y = 172.4 + 6.4(x - 68.8) \quad \text{or}$$

$$y = -267.92 + 6.4x$$

Based on this equation, someone who is 72 inches tall is estimated to weigh 192.88 pounds.

*8. The Standard Error of the Estimate, denoted $S_{yx}$, is a method used to indicate how well the estimated regression line fits the data. It measures the average difference between the observed and predicted values.

$$S_{yx} = \sqrt{\frac{\Sigma(y - y_p)^2}{n - 2}}, \qquad \text{where}$$

$y_p$ = predicted value, y = observed value and n = the number of paired observations.

Example: For each of the given heights in the above example $y_p$ must be found. Then $\Sigma(y - y_p)^2$ can be determined.

| x | y | $y_p$ | $y - y_p$ | $(y - y_p)^2$ |
|---|---|---|---|---|
| 68 | 170 | 167.28 | +2.72 | 7.40 |
| 66 | 161 | 154.48 | +6.52 | 42.51 |
| 71 | 190 | 186.48 | +3.52 | 12.39 |
| 69 | 172 | 173.68 | −1.68 | 2.82 |
| 69 | 155 | 173.68 | −18.68 | 348.94 |
| 68 | 180 | 167.28 | +12.72 | 161.80 |
| 65 | 143 | 148.08 | −5.08 | 25.81 |
| 72 | 200 | 192.88 | +7.12 | 50.69 |
| 67 | 158 | 160.88 | −2.88 | 8.29 |
| 73 | 195 | 199.28 | −4.28 | 18.32 |
| | | | TOTAL | 678.97 |

$$S_{yx} = \sqrt{\frac{678.97}{8}}$$

$$S_{yx} \approx 9.2$$

B. Performance Objectives

1. The student can sketch the estimated regression line in a scatter diagram.

2. The student can determine the equation for the estimated regression line.

3. The student can use the least squares method to predict the dependent variable's value.

4. The student understands that the observed value will vary from the predicted value.

5. The student understands that when the correlation between two variables is low the differences between observed and predicted values will generally be quite high.

6. The student understands that when the correlation between two variables is high, observed and predicted values are expected to be close.

7. The student understands that "b" in the regression equation is the slope of the line.

8. The student understands that the sign of the correlation coefficient indicates the direction of slope for the regression line.

9. The student understands the role of correlations in making predictions.

10. The student understands the algebraic equation for a straight line.

11. The student can plot a straight line on a graph given an equation.

*12. The student understands that the standard error of the estimate indicates how well the least squares prediction equation (equation of the regression line) describes the relationship between two variables.

*13. The student understands that to find $S_{yx}$ the regression equation must first be used on all observed values of x to find the predicted values of y.

*14. The student understands that the smaller $S_{yx}$ is, the more precise the future predictions are likely to be.

*15. The student understands that $S_{yx} = 0$ means the least squares regression equation is a perfect description of the relation between x and y.

*16. The student can compute the standard error of the estimate.

*17. The student understands that roughly 68% of the time the actual value of the dependent variable will not differ from the predicted value by more than $S_{yx}$.

*18. The student understands that roughly 95% of the time the actual value of the dependent variable will not differ from the predicted value by more than $2(S_{yx})$.

C. Sample Exercises and Solutions

1. (P.O. 9) What is the purpose of the regression equation?

   *Solution:*
   *It is used to make predictions about the value of the dependent variable.*

*2. (P.O. 15) True or False: The standard error of the estimate equals 0 when r = -1.

   *Solution:*
   *True. $S_{yx} = 0$ means that the regression equation describes the relation perfectly. In other words the ordered pairs of observations lie on a straight line. $r = -1$ also means there is perfect correlation and the points lie on a straight line.*

3. (P.O. 1,2,3) Eight junior high school students received the following grades on their English and foreign language exams:

| Student | English Score | Foreign Language Score |
|---|---|---|
| Betty | 69 | 57 |
| Lukas | 81 | 90 |
| Kimo | 53 | 55 |
| Maile | 93 | 88 |
| LiAnn | 79 | 75 |
| David | 83 | 83 |
| Gloria | 86 | 81 |
| Jeff | 75 | 69 |

   a. Construct a scatter diagram.

   b. Compute the regression equation to predict a student's foreign language score.

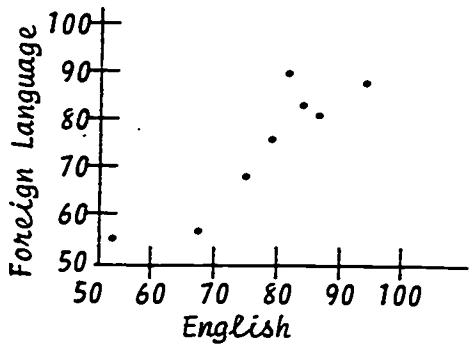
   c. If a student receives a 71 in English, what would you expect him/her to receive in foreign language?

   d. Compute the regression equation to predict a student's English score.

e. If a student receives an 87 in foreign language, what would you expect him/her to receive in English?

*Solution:*

a.




b. *The independent variable, x, is English score and y, the dependent variable, is Foreign Language score.* $\Sigma x = 619$, $\Sigma x^2 = 48{,}931$, $\Sigma y = 598$, *and* $\Sigma xy = 47{,}277$

$$b = \frac{8(47{,}277) - 619(598)}{8(48{,}931) = (619)^2} = \frac{8{,}054}{8{,}287} \approx 0.97$$

$$\bar{x} = \frac{619}{8} = 77.375 \qquad \bar{y} = \frac{598}{8} = 74.75$$

*Therefore,*

$$y = 74.75 + 0.97(x - 77.375) \quad \textit{or}$$

$$y = -0.304 + 0.97x$$

c. $y = -0.304 + 0.97(71) = 68.566$

d. *In this case the Foreign Language Score is x, the independent variable, and the dependent variable, y, is English.* $\Sigma x = 598$, $\Sigma y = 619$, $\Sigma xy = 47{,}277$, *and* $\Sigma x^2 = 45{,}954$

$$b = \frac{8(47{,}277) - 598(619)}{8(45{,}954) - (598)^2} = \frac{8{,}054}{10{,}028} \approx 0.80$$

$$\bar{x} = \frac{598}{8} = 74.75 \qquad \bar{y} = \frac{619}{8} = 77.375$$

*Therefore,*

$$y = 77.375 + 0.8(x - 74.75) \quad \textit{or}$$

$$y = 17.575 + 0.8x$$

e. $y = 17.575 + 0.8(87) = 87.175$

4. (P.O. 13,14,16,17) Use the data in Exercise 3 to answer the following:

a. Find the standard error of the estimate for the dependent variable, Foreign Language.

b. If a student scored 84 on the English test, 68% of the time the student's Foreign Language score would be between which two numbers?

*Solution:*

a.

| $x$ | $y$ | $y_p$* | $y - y_p$ | $(y - y_p)^2$ |
|---|---|---|---|---|
| 69 | 57 | 66.626 | -9.626 | 92.7 |
| 81 | 90 | 78.266 | +11.734 | 137.7 |
| 53 | 55 | 51.106 | +3.894 | 15.2 |
| 93 | 88 | 89.906 | -1.906 | 3.6 |
| 79 | 75 | 76.326 | -1.326 | 1.8 |
| 83 | 83 | 80.206 | +2.794 | 7.8 |
| 86 | 81 | 83.116 | -2.116 | 4.5 |
| 75 | 69 | 72.446 | -3.446 | 11.9 |
| | | | TOTAL | 275.2 |

$$S_{yx} = \sqrt{\frac{275.2}{6}}$$

$$S_{yx} \approx 6.8$$

*The equation $y = -0.304 + 0.97x$ is used here.

b. $y = -0.304 + 0.97(84) = 81.176$
The student would be expected to score between 74.376 and 87.976 68% of the time $(81.176 - 6.8 = 74.376$ and $81.176 + 6.8 = 87.976)$.

5. The following problem set from Statistics By Example is also an appropriate exercise for use with this objective:

"Points and Fouls in Basketball" by Albert P. Shulte in Exploring Data.

D. Suggested Student Readings

1. Battan, Louis J., "Cloud Seeding and Rainmaking" in Statistics: A Guide to the Unknown.

2. Campbell, Donald T., "Measuring the Effects of Social Innovations by Means of Time Series" in Statistics: A Guide to the Unknown.

3. Cliff, Norman, "Adverbs Multiply Adjectives" in Statistics: A Guide to the Unknown.

4. Miller, Robert G., "The Probability of Rain" in Statistics: A Guide to the Unknown.

5. Street, Elisabeth and Mavis G. Carroll, "Preliminary Evaluations of a New Food Product" in Statistics: A Guide to the Unknown.

6. Tufte, Edward R., "Registration and Voting" in Statistics: A Guide to the Unknown.

## AVAILABLE MATERIALS

Until recently statistics has been taught primarily to college students and highly motivated high school seniors. Consequently the number of available materials for teaching an option y course is very limited. The textbooks tend to be aimed at college level students. However, more and more schools are adding lower-level statistics courses to their curriculum and the need for more appropriate materials has been recognized. Development is underway, but it does take time. Meanwhile, there are a few materials aimed at high school students and there are a few college texts that can be adapted for use in the classroom. In addition to forthcoming texts, the National Council of Teachers of Mathematics 1981 yearbook on Statistics and Probability should provide some interesting and worthwhile suggestions for teaching this course.

TEXTS (ANNOTATED)

Freedman, David, Robert Pisani, and Roger Purves

STATISTICS
W. W. Norton & Co., Inc., 1978

The authors of this text believe that too much mathematical notation can be confusing and that too often not enough emphasis is placed on understanding concepts. Their approach emphasizes the use of models and real-life illustrations and applications.

Part II. Descriptive Statistics and Part III. Correlation and Regression would provide the major content for an option y Statistics course.

Chapter 3 (Part II) develops the concept and understanding of histograms very well. The histogram is then used throughout the text to develop and relate many other concepts. Other graphing techniques such as bar graphs and circle graphs are not introduced. The construction of frequency distributions is also omitted. Chapters on the mean, standard deviation, correlation and regression are presented very well and include many excellent examples.

In general, the book uses a narrative format with everyday language to describe statistical ideas. Mathematical formulas are introduced only after basic concepts have been developed. There are many exercises throughout the text. Some of them provide practice in computational skills while others cannot be solved by plugging numbers into formulas. They involve estimates and qualitative judgments so that students can come to grips with the concepts.

This book makes an excellent addition to the classroom library even if not adopted as a text. It will provide the teacher with many ideas for presentation of material as well as possible test and quiz items. It will also benefit the student who needs help in understanding a concept.

| | Code | Price |
|---|---|---|
| Hardcover, 589 pages | 0-393-09076-0 | $13.95 |
| Instructor's Manual | 0-393-09041-8 | ---- |

Freund, John E.

STATISTICS, A FIRST COURSE
Prentice-Hall, Inc., 1976

This textbook is easy to read and includes some very good explanations and examples of new concepts. All of the learner objectives are covered. However, the organization of the text requires some work with probability distributions and normal distributions before variance, standard deviation and average deviation are explained. To include this work in the course *may* cause difficulty time-wise in covering the remaining objectives.

Some topics are covered a little too quickly. In some instances new vocabulary is introduced and developed in the exercises. This allows the better students to extend their background and gives some flexibility to the instructor. Knowledge of these concepts is not required in future chapters.

Supplementary materials designed to complement the text are included in a student workbook. The workbook chapters correspond to the chapters in the text and include a Glossary, Formulas, True-False, Multiple Choice, and Completion Items. Review sections are also included.

| | Code | Price |
|---|---|---|
| Cloth, 374 pages | 84605-5 | $11.16 |
| Student Workbook | 84601-4 | 4.40 |

Ingram, John A.

ELEMENTARY STATISTICS
Addison-Wesley Publishing Co., 1977

The author states that the "theme of this work is using good judgment and exercising a sense of reasonableness." The book is written for beginners, uses clear language and is fairly easy to read. An appendix on Math Essentials is included for students who need the review.

All of the learner objectives for the statistics course are covered although the desired depth is not always there. Explanations tend to be a little too brusque for this level of student. Although they are used to illustrate correlation and regression examples, scattergrams are never defined. Knowledge of how to construct one is assumed. Also, linear regression is discussed before correlation coefficients. The concepts of slope, y-intercept, and linearity must be understood before the section on regression can be mastered.

Supplementary materials designed to complement the text are included in a student workbook. The material for each chapter includes Key Words, a chapter Overview, Procedure Guides, Self-Study Questions and Study Problems. A data base of characteristics of 234 students is used throughout the workbook.

This is a usable text, but the workbook and/or a lot of teacher support are necessary for this level student.

| | Code | Price |
|---|---|---|
| Hardcover, 445 pages | 52660 | $12.76 |
| Instructor's Guide | 52662 | 0.80 |
| Student Workbook | 52661 | 3.96 |

Newmark, Joseph

STATISTICS AND PROBABILITY IN MODERN LIFE
Holt, Rinehart and Winston, 1977

This text is written in clear everyday language that the student should be able to understand. The amount of new terminology introduced is held to a minimum. Each chapter begins with newspaper or magazine clippings illustrating the concepts discussed in the chapter.

All of the learner objectives for the statistics course are covered. Because this is a non-mathematical text the rigor of a mathematical explanation is often omitted. In general, this works well for students at this level. However, there are times when not enough of an explanation is given for the student to thoroughly understand the concept. Many examples are provided throughout the chapter and exercises are plentiful.

Depending upon the mathematical rigor desired in this course, the teacher may want to supplement explanations and/or to introduce additional terminology.

| | Code | Price |
|---|---|---|
| Cloth, 516 pages | 03-018881-4 | $14.95 |
| Solution Manual | 03-022216-8 | 3.20 |

Schmidt, Marty J.

UNDERSTANDING AND USING STATISTICS *BASIC CONCEPTS*
D. C. Heath and Company, 1975

This textbook emphasizes "principles and ideas that are the basis of applied statistics rather than including a large number of different methods." It is intended as either a primary text or as a supplement to provide the student with a clear conceptual framework from which to build.

Part I/Descriptive Statistics would provide the major content for an option y Statistics course. The minimum course would have to include the Introduction (which describes the nature of statistics), Chapter 2: The Frequency Distribution, Chapter 3: Central Tendency, Chapter 4: Variability, Chapter 6: Correlation and Chapter 7: Regression.

The book uses a narrative format and attempts to promote understanding of concepts. Many conceptual examples and fewer numerical examples are provided.

Because of the narrative format and the language of statistics, the reading level is higher than many other textbooks. Also, the appropriate chapters include enough material in addition to the required basics that a teacher using the book must be selective. Otherwise, it would be difficult to complete the core course during the semester. The exercises at the end of each chapter are good, but very limited in number.

Even if this book is not adopted as a text, it makes a very good reference book for both teachers and students.

| | Code | Price |
|---|---|---|
| Paperback, 361 pages | 0669-94490-4 | $ 7.95 |

Campbell, Stephen K.

FLAWS AND FALLACIES IN STATISTICAL THINKING
Prentice-Hall, Inc., 1974

The purpose of this non-technical book is to help lay persons to "increase their ability to judge the quality of statistical evidence and, in turn, to make better informed decisions of many facets of everyday life."

The sequence of topics is similar to that found in most beginning texts. Chapters include: Dangers of Statistical Ignorance, Meaningless Statistics, Cheating Charts, Accommodating Averages, Improper Comparisons and Relationships: Causal and Casual. The terminology and manner of presentation assume no previous knowledge on the part of the reader, although occasional terms such as logarithm may cause difficulty.

This book has good examples and makes interesting reading.

| | Code | Price |
|---|---|---|
| Paperback, 220 pages | 013-32214-4 | $6.95 |

Huff, Darrell

HOW TO LIE WITH STATISTICS
W. W. Norton and Company, Inc., 1954

This non-technical book illustrates how statistics can be used to demonstrate any point of view. It is a "primer in ways to use statistics to deceive." Although the book is over 25 years old, the examples are still relevant.

Chapters in the book include: The Well-Chosen Average, The Gee-Whiz Graph, and How to Talk Back to Statistics. Most examples are humorous and make interesting reading. The terminology and manner of presentation assume no previous knowledge on the part of the reader.

This supplementary book makes good reading for both teacher and student and is a very good addition to the classroom library.

| | Code | Price |
|---|---|---|
| Paperback, 142 pages | 0-393-09426-X | $1.95 |

STATISTICS & INFORMATION ORGANIZATION
Creative Publications, 1978

This resource book is divided into five major sections: Content for Teachers, Didactics, Teaching Emphasis, Classroom Materials, and Glossary and Annotated Bibliography. The Classroom Materials section has two major subsections: Concepts and Skills and Applications. All five sections are further subdivided by topic.

The book is meant to be a topical resource for teachers and the classroom materials are intended to supplement adopted programs or textbooks.

There are over 450 pages of student materials for worksheets, calculator activities, games, puzzles, bulletin board suggestions, project ideas and so on. Reproduction permission is granted. Headings appropriate to this course include Tables, Graphs, Scatter Diagrams, Misleading Statistics, Mean, Median, Mode and Range & Deviation.

Although the resource is intended for use in grades 5-9, much of the material is appropriate for older students as well. Ideas from some of the remaining material can be adapted by the classroom teacher.

| | Code | Price |
|---|---|---|
| Ready to put in binder; 888 pages | 10684 | $27.50 |

MATHEMATICS IN SCIENCE AND SOCIETY
Creative Publications, 1978

This book deals with mathematics as it relates to astronomy, biology, environment, music, physics, and sports. Each of these topics is described with an introduction, classroom materials, glossary, and annotated bibliography. Many of the materials are appropriate to the study of statistics.

| | Code | Price |
|---|---|---|
| Ready to put in binder; 464 pages | [illegible]83 | $22.50 |

Mosteller, Frederick, William Kruskal, Richard F. Link, Richard S. Pieters, and Gerald R. Rising, editors

STATISTICS BY EXAMPLE
Addison-Wesley Publishing, 1973

This is a series of four pamphlets providing real-life problems in probability and statistics for the secondary school level. They are Exploring Data, Weighing Chances, Detecting Patterns, and Finding Models. Each pamphlet is intended to stand alone, but cross-referencing is given. The materials were prepared by the Joint Committee on the Curriculum in Statistics and Probability of the American Statistical Association and the National Council of Teachers of Mathematics.

Exploring Data "shows how to organize data tabularly and graphically to get messages and to give them, and it introduces elementary probability in circumstances where simple counting gets one off to a good start."

Weighing Chances "develops probability methods through random numbers, simulations, and simple probability models, and presents serious analyses of complex data using informally the idea of scatter and residuals."

Detecting Patterns "presents several standard statistical devices--the normal distribution, the chi-square test, and regression methods."

Finding Models "encourages the student to develop models as structures for data, so that departures from models can be recognized and new structures built."

Each problem is explained in easy-to-understand language and includes exercises and projects relative to the example. They can be used by an entire class, for enrichment or for extra-credit work. The pamphlets are excellent supplements to this course with the first two, Exploring Data and Weighing Chances, being especially appropriate.

| | Code | Price |
|---|---|---|
| Exploring Data, 125 pages | 04873 | $3.48 |
| Teachers' Commentary and Solutions Manual | 04874 | 2.13 |
| Weighing Chances, 145 pages | 04875 | 3.48 |
| Teachers' Commentary and Solutions Manual | 04876 | 2.13 |
| Detecting Patterns, 166 pages | 04877 | 3.48 |
| Teachers' Commentary and Solutions Manual | 04878 | 2.13 |
| Finding Models, 146 pages | 04879 | 3.48 |
| Teachers' Commentary and Solutions Manual | 04880 | 2.13 |

Tanur, Judith M. et. al., editors

STATISTICS: A GUIDE TO THE UNKNOWN
Holden-Day, Inc., 1972

This book contains a series of 44 essays on the applications of statistics. It was prepared by the Joint Committee on the Curriculum in Statistics and Probability of the American Statistical Association and the National Council of Teachers of Mathematics.

"The essays explain statistical ideas and contributions without dwelling on their mathematical aspects." They illustrate a wide variety of fields of applications. The authors were asked "to stress one or a very few important problems within their field of application and to explain how statistics and probability help to solve them and why the solutions are useful to the nation, to science, or to the people who originally posed the problem."

This book makes an excellent supplement to any text.

| | Price |
|---|---|
| Paperback, 430 pages (available through NCTM) | $7.50 |

# TEACHING STATISTICS

"Teaching Statistics is a new journal (Spring, 1979) for all who use statistics in their teaching of pupils and students aged 9 to 19. It is sponsored by the Applied Probability Trust, the Institute of Statisticians, the International Institute, and the Royal Statistical Society. Teaching Statistics seeks to inform, enlighten, entertain and encourage all teachers to make better use of statistical ideas."

Contents include "articles on teaching statistics, statistics in mathematical lessons, ... ideas for project work, basic concepts in statistics, ... ideas for classroom teaching, historical developments, ... etc." Also included are "authoritative reviews of books and equipment, correspondence and problems solved."

Three issues are to be published each year. Subscriptions are $6.00 and should be sent to:

Peter Holmes, Editor
Department of Probability and Statistics
The University
Sheffield S10 2TN
UNITED KINGDOM

ADDITIONAL RESOURCES

The following is a list of books and/or pamphlets which can be used as teacher references, student references, and/or student supplements.

1. Applied Statistics, Alfred D. Garvin. Portland, Maine: J. Weston Walch, Publisher, 1977. (124 pages, paperback, teacher/student reference).

2. "Careers in Statistics." Washington, D.C.: American Statistical Association.

   Pamphlet describes occupations in area of statistics. A variety of applications are described and information on higher education in statistics is given.

3. Collecting, Organizing, and Interpreting Data, Topics in Mathematics, Booklet Number 16. Reston, Virginia: National Council of Teachers of Mathematics, 1969. (42 pages, paperback, teacher reference).

4. The Figure Finaglers, Robert S. Reichard. New York: McGraw-Hill Book Co., 1974. (274 pages, paperback, teacher reference, student supplement).

5. Guidelines for Teaching Mathematics, Donovan A. Johnson and Gerald R. Rising. Belmont, California: Wadsworth Publishing Co., Inc., 1967. (438 pages, hardcover, teacher reference).

6. How to Take a Chance, Darrell Huff. New York: W. W. Norton and Company, Inc., 1959. (173 pages, paperback, teacher/student reference, student supplement).

   Although the book emphasizes probability more than statistics, there are sections that are appropriate to this course.

7. Introduction to Statistical Analysis and Inference, Sidney J. Armore. New York: John Wiley and Sons, Inc., 1966. (546 pages, hardcover, textbook, teacher/student reference).

   This is an excellent teacher reference for course content. Complete, detailed coverage of topics is provided. Explanations are very good. Understanding of material is emphasized.

8. Making Mathematics, Books 1-4, D. Paling, C. S. Banwell, K. D. Saunders. New York: Oxford University Press, 1972. (125 pages each, booklets, teacher reference).

   These booklets cover many topics and are intended for grades 4-8. However, activities for statistics are included and can be adapted for use in this course by the teacher.

9. Mathematics: A Human Endeavor, Harold Jacobs. San Francisco: W. H. Freeman and Co., Publishers, 1970. (529 pages, hardcover, teacher/student reference, student supplement).

Includes an interesting chapter on statistics with some very good examples and exercises.

10. Overview and Analysis of School Mathematics, Grades K-12. Washington, D.C.: Conference Board of the Mathematical Sciences, 1975.

11. Probability and Statistics, Stephen J. Willoughby and Bruce R. Vogeli. Morristown, New Jersey: Silver Burdett Company, 1968. (216 pages, hard-cover, teacher reference).

Contains good chapter on organizing and reporting data.

12. Readings in Mathematics, Book 2, Irving Adler, editor. Lexington, Massachusetts: Ginn and Company, 1972. (188 pages, paperback, teacher reference, student supplement).

13. Sampling and Statistics, Applications in Mathematics, Course A, Donovan A. Johnson, et al. Glenview, Illinois: Scott Foresman and Co., 1972. (62 pages, booklet, teacher reference).

This activity-oriented booklet is intended as a text, but it is too elementary for this course. However, there are some good examples for the teacher to draw upon.

14. Schools Council Project on Statistical Education, project papers. United Kingdom: Schools Council Project on Statistical Education, 1976.

Series of ten papers on the present position of statistics in the secondary schools.

15. Science in Geography, Books 1-4. New York: Oxford University Press, 1974. (Booklets, teacher reference).

Set of four books which describe a scientific approach to the study of geography. Many examples and problems presented which require the use of statistical methods.

16. Statistical Thinking: A Structural Approach, John L. Phillips, Jr. San Francisco: W. H. Freeman and Co., Publishers, 1973. (124 pages, paperback, teacher reference, student supplement, high reading level).

17. "Statistics," Richard S. Pieters and John J. Kinsella. The Growth of Mathematical Ideas, Grades K-12. Reston, Virginia: National Council of Teachers of Mathematics.

Provides an excellent overview of statistics at the pre-college level and makes recommendations for the teaching of statistics in grades K-12.

18. Statistics and Probability, S. E. Hodge and M. L. Seed. London: Blackie and Son, Limited and W. and R. Chambers, Limited, 1972. (264 pages, paperback, teacher reference, high reading level).

19. Statistics in the Real World, A Book of Examples, Richard J. Larsen and Donna Fox Stroup. New York: Macmillan Publishing Co., Inc., 1976. (245 pages, paperback, teacher reference).

This is a workbook "... designed to show how statistical procedures can be applied to real problems." Many of the examples include advanced topics, but others cover some descriptive statistics and regression and correlation. The book provides the instructor with some appropriate homework problems, class exercises and review materials.

20. Statistics Made Relevant: A Casebook of Real Life Examples, Paul Brown and Ernest M. Scheuer. New York: John Wiley and Sons, Inc., 1976. (217 pages, paperback, student supplement).

This workbook contains 48 cases drawn largely from the news media. Many of the cases, however, involve more advanced concepts.

21. The Teaching of Probability and Statistics, Lennart Råde, editor. New York: John Wiley and Sons, Inc., 1970. (373 pages, hardbound, teacher reference, high reading level).

This is a collection of papers presented at the first CSMP International Conference on the Teaching of Probability and Statistics. It contains many ideas for classroom activities although some ideas are very advanced.

22. Theory and Problems of Statistics, Schaum's Outline Series, Murray R. Spiegal. New York: McGraw-Hill Book Company, 1961. (359 pages, paperback, teacher reference).

This book contains 875 problems solved in detail.

23. The Use and Misuse of Statistics, David Newton. Portland, Maine: J. Weston Walch, Publisher, 1973. (42 pages, paperback, teacher reference, student supplement).

SOURCES OF DATA


1. "Annual Economic Review." Bank of Hawaii, Department of Business Research, Financial Plaza of the Pacific, Honolulu, Hawaii 96813. (FREE)

2. Consumer Reports, the (year) Buying Guide Issue. Consumers Union of the United States.

3. County of Hawaii Data Book. Department of Research and Development, 25 Aupuni Street, Hilo, Hawaii 96720. (FREE)

4. "Economic Indicators." A monthly report by First Hawaiian Bank, Research and Planning Division, 156 S. King Street, Suite 1505, Honolulu, Hawaii 96813. (FREE)

5. Guinness Book of World Records, by Norris McWhirter and Ross McWhirter.

6. "Hawaii County--Facts and Figures," The Big Island Report. Annual pamphlet by Department of Research and Development, County of Hawaii, 25 Aupuni Street, Hilo, Hawaii 96720. (FREE)

7. Historical Statistics of Hawaii, by Robert C. Schmidt. University Press of Hawaii, 1977.

8. Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition. U.S. Bureau of the Census.

9. Information Please Almanac. Dan Golen Paul Associates.

10. The Official Associated Press Sports Almanac. The Associated Press.

11. Pocket Data Book. U.S. Bureau of the Census. (For sale by Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 or any U.S. Department of Commerce Field Office.)

12. Reader's Digest (year) Almanac and Yearbook. Reader's Digest Association, Inc.

13. Research and Statistics Offices of most city, state and federal agencies publish pamphlets and brochures containing statistics relevant to their area.

14. The State of Hawaii Data Book: A Statistical Abstract. An annual publication of the Department of Planning and Economic Development, Information Office, 7th Floor, Kaimalu Building, 250 S. King Street, Honolulu, Hawaii 96804. ($4.00 locally and $5.00 post-paid out-of-state)

15. Statistical Abstract of the United States. U.S. Bureau of the Census. (Also published as The U.S. Fact Book by Grosset and Dunlop.)

16. Webster's New Collegiate Dictionary. G. & C. Merriam Company.

17. The World Almanac and Book of Facts. Newspaper Enterprise Association.

18. The World in Figures, by Victor Showers.

19. World Statistics in Brief. United Nations.

83

Films Incorporated:

Statistics (16-mm Sound B/W Film; No. 324-0058; 13 min.; $180 sale; $25 rental)

"... Bar graphs and circle graphs may be used to present statistics. The film illustrates problems relating to world food supply and world population distribution with graphs. Surprising facts emerge and lead to the study of related statistics. ...
... the film stresses that we must be sure that our statistics tell the true story and the whole story. For example, there are many more auto accidents in 1975 than in 1939. Are today's drivers worse or are there other contributing causes? Many interesting examples leading to ridiculous conclusions are given to illustrate the point that, 'Figures never lie but liars often figure.' "

Film Services of Hawaii:

Fact Finder for the Nation (Census Bureau/U.S. Dept. of Commerce; 29 min.; color; #30472)

"The Census Bureau is the only agency of the government that comes in contact with every person in the country. The film shows how the census is taken every ten years, as well as the wide scope of work the Bureau does during the other nine years--work which provides statistics about how the people and the economy change on a weekly, monthly, quarterly, and annual basis."

We (Census Bureau/U.S. Dept. of Commerce; 29 min.; color; #30473)

"This film, narrated by Burgess Meredith, presents key findings of the 1970 census--the enormous growth of our suburbs, the continued movement of the population from rural to urban areas. The statistics about the progress and problems of America are put into human terms, concentrating on the people as they work, play and relax."

(All films are 16-mm sound films available to educators on a FREE-loan basis.)

Indiana University Audio Visual Center:

Statistics at a Glance (16-mm, sound; 27 min.; color; ISC-725; $12.50 rental)

"Explains, using computer graphics, that descriptive statistics is a method for organizing and summarizing data so that the data can be understood quickly and expressed in precise, accurate terms. Illustrates the use of measures of central tendency, such as mean, median, and mode in a frequency distribution. Demonstrates the use of variability, standard deviation, Z-scores, percentile rank, and correlation as they relate to frequency distribution."

Inferential Statistics: Sampling and Estimation Saves $122,000,000!
(16-mm, sound; 20 min.; color; ISC-769; $11.50 rental)

"Presents a situation in which the use of statistics saves a company from promoting a product that the public would not buy to introduce basic concepts in statistics. Explains the concepts and applications of defining a population, selecting a random sample, using tables of random numbers, and conducting surveys. Demonstrates the inference of interval estimates of the population from the statistics of the sample. Concludes with a familiar, amusing twist."

Michigan State University:

The Fact Finders (16-mm, sound; 31 min.; color; $5.00 rental)

"How farm statistics are collected and used ... from the farmer through the state's FEDERAL CROP REPORTING SERVICE back to farms and agribusiness. Shows crucial national crop and livestock reports being prepared and released under rigid security regulations."

Statistics at a Glance

See listing under Indiana University.

Food Processors Institute:

USING SQC (46-2x2 color slides; 17 minute recorded tape; printed-illustrated script; reference material. With cassette tape - FA0510 ... $55.00; with reel-to-reel tape - FA0520 ... $55.00)

"Explains the use, value, and meaning of statistical quality control carts in a non-technical manner and in simple mathematical terms. Filling operations and labeling are used as examples."

Great Plains National:

MAINLY MATH, PROGRAM 14, STATISTICS AND GRAPHS (videotape or video cassette; 20 min.; $25.00 rental; $175.00 sale)

"The opening vignette of an executive board meeting emphasizes the ever increasing problem of properly communicating information. This lesson focuses on the topics of statistics and graphs--tools used in organizing and classifying data ... to graphically communicate information."

MAINLY MATH, PROGRAM 15, STATISTICS AND GRAPHS II (Videotape or video cassette; 20 min.; $25.00 rental; $175.00 sale)

"A group of students decide to construct a graph. The lesson shows how to organize and construct understandable graphs--whether they be of the bar, parts-whole or line variety."

Lansford Publishing Co.:

STATISTICS FOR THE NON-STATISTICIAN (2 cassettes; 1162AD ... $35.95)

"Provides the non-statistician with an overview of descriptive statistics and an overview of inferential statistics."

PRESENTATION OF DATA (40 slides; teacher guide; 1 cassette; 1240AD ... $99.95)

"Covers the following subjects: use of graphic presentations, presenting data in tabular form, pie charts, simple vertical and horizontal bar charts, multiple bar charts, simple and multiple line charts, statistical maps, pictographs, logarithmic scales, and related concepts of charting. Also covers methods to improve graphic presentations, applications of basic charting methods, and how to avoid creating misleading impressions in graphic presentations."

DESCRIPTIVE STATISTICS (40 slides; teacher guide; workbook; 1 cassette; 1241AD ... $109.95)

"Deals with frequency distributions; measures of central tendency, and measures of dispersion. Subjects covered include: the nature of descriptive statistics; arrays; tallies; frequency distribution tables; frequency polygons; histograms; shapes of frequency distributions; discrete vs. continuous distributions; the normal distribution; mean; median; mode; range, mean deviation; variance and standard deviation. Treatment of these subjects is at an introductory level with emphasis on understanding of concepts, and examples and applications."

Prismatron Productions, Inc.:

STATISTICAL DIAGRAMS (45 slides; cassette; program guide; S1 ... $85)
"Large tables of statistical data can often be presented more concisely and meaningfully by a suitable diagram, but care must be taken to insure that complete accuracy is maintained. This program provides an introduction to the representation of statistical data by diagrams, and emphasis is laid throughout on the correct use of scales and keys to guard against misinterpretation. Following a discussion of the construction of graphs, the program describes how the results of a smoking survey might be displayed by histograms or pictograms. A new example forms the basis of an explanation of several kinds of bar charts, and the program concludes by describing the steps involved in drawing a pie diagram."

WHAT IS 'AVERAGE'? (40 slides; cassette; program guide, S2 ... $85)
"Many people do not realize that when they talk glibly of a team's goal average or the average temperature of a holiday, they are actually referring to the arithmetic mean. However, this is not always the most appropriate measure to use. This program explores the differences between the arithmetic, geometric and harmonic means, showing how to calculate them and when to use them. The arithmetic mean is also compared with the mode and median of several types of probability distribution. This is an elementary but extremely informative program."

SAMPLING AND QUESTIONNAIRES (44 slides; cassette; program guide; S3 ... $85)
"Modern advertising and opinion polls mean that everyone is familiar with the concept of taking samples and administering questionnaires. This program provides a lighthearted introduction to the problems of sample selection and questionnaire design. The cartoon-style diagrams enable many important points to be made extremely simple and clear. The merits of collecting information by direct interview and postal questionnaire are compared, and the basic principles underlying the construction of a good questionnaire are discussed. The necessity of choosing a valid sampling technique is stressed, and the concepts of random and stratified sampling are described in more detail. Illustrative examples are drawn from a wide variety of fields."

DISTRIBUTIONS (46 slides; cassette; program guide; S6 ... $95)
"Distributions play a fundamentally important part in the study of Statistics, and students need to acquire an understanding of the relationship between a distribution curve and the data it represents. This program makes use of many different examples of discrete and continuous variables to explain the construction of line diagrams, frequency polygons and distribution curves. The calculation of the mean and various measures of spread is explained, and the meaning of terms such as skew and kurtosis is discussed. The calculation and use of the mode, median and fractile measures are described."

THE NORMAL DISTRIBUTION (40 slides; cassette; program guide; S7 ... $85)
"Since this is the most important distribution, it is essential for the student to have a thorough comprehension of what is meant by a normally distributed variable. This is a quick-paced, highly visual program in which the Normal distribution curve is derived as an approximation to a histogram representation of a distribution of mean's heights. The significance of the two descriptive parameters, the mean and the standard deviation, is explored in depth, and a discussion of the area under the curve leads naturally to an explanation of converting any Normal distribution to the Standard Normal. The far-reaching importance of this distribution is illustrated by examples taken from quality control, medical research and other fields."

## TRANSPARENCIES

| GAF Corporation: | Code | Price |
|---|---|---|
| Introduction to Statistics (complete set, 29 transparencies) | 253-760 | $144.00 |
| Concepts (19 transparencies) | 253-761 | 96.00 |
| What is Statistics? (2) | 259-489 | 6.40 |
| What do Statisticians Do? (3) | 259-490 | 7.70 |
| Graphic Presentations (2) | 259-491 | 5.25 |
| Bar Graph (2) | 259-492 | 5.25 |
| Pie Graph (1) | 259-493 | 5.50 |
| Pictograph (2) | 259-494 | 5.25 |
| Map Chart | 259-495 | 4.05 |
| Frequency Distribution Table & Line Graph (3) | 259-496 | 6.60 |
| Grouped Frequency Distribution Table, Histogram & Frequency Polygon (4) | 259-497 | 7.75 |
| The Average Measures of Central Tendency (2) | 259-498 | 5.25 |
| The Arithmetic Mean (2) | 259-499 | 5.25 |
| The Median (2) | 259-503 | 5.25 |
| The Mode (1) | 259-504 | 4.10 |
| Measures of Variability-Range (2) | 259-505 | 5.25 |
| The Normal Frequency Distribution Curve (2) | 259-509 | 5.25 |
| Relationship between Variables, Correlation & Regression (2) | 259-512 | 5.25 |
| Regression-Scatter Diagram (1) | 259-513 | 4.10 |
| Sampling (2) | 259-516 | 5.25 |
| Sampling Methods (2) | 259-517 | 5.25 |
| Calculations (10 transparencies) | 253-762 | 48.00 |
| Calculation of the Arithmetic Mean from a Grouped Frequency Distribution I (1) | 259-500 | 4.10 |
| Calculation of the Arithmetic Mean from a Grouped Frequency Distribution II (1) | 259-501 | 4.10 |
| Calculation of the Aritmetic Mean from a Grouped Frequency Distribution by the Class Interval Method (3) | 259-502 | 6.60 |
| The Standard Deviation (2) | 259-506 | 5.25 |
| Calculation of Standard Deviation from Grouped Data (1) | 259-507 | 4.10 |
| Calculation of Standard Deviation from Group Data by Class Interval Method (1) | 259-508 | 4.10 |
| Characteristics of the Normal Curve (1) | 259-510 | 4.10 |
| Use of the Normal Curve (2) | 259-511 | 5.65 |
| Regression-Calculation of Best Fitting Straight Line (1) | 259-514 | 7.35 |
| Correlation-Calculating the Degree of Association Between Two Variables (1) | 259-515 | 4.10 |

Lansford Publishing Co.:

EFFECTIVE PRESENTATION OF DATA
"Deals with effective presentation of data in graphic forms, such as pie charts, bar charts, simple line charts, multiple bar and line charts, pictographs, statistical maps, and others. Fundamental applications and use of logarithmic scale and charts are discussed."
W215TAD (17 transparencies, workbook, teacher guide) ............ $159.95

FREQUENCY DISTRIBUTIONS
"Includes the following: measurements and scales; graphic presentations; arrays and frequency distribution tables; frequency polygons; histograms; cumulative distributions; discrete versus continuous distributions; shapes of different distributions; and percentiles."
W301TAD (14 transparencies, workbook, teacher guide) ............ $139.95

HOW TO CONDUCT A SURVEY
"A brief course for classes where non-statisticians are expected to know something about survey research. Covers problem definition, questionnaire design, sampling schemes, tabulation, and presentation of data. Very practical."
568AD (17 transparencies, teacher guide, sample size calculator, booklets) ................................................... $149.95

QUESTIONNAIRE CONSTRUCTION
"Enables the teacher or trainer to explain how a good questionnaire is designed, how to avoid common pitfalls in question wording, the variety of answer formats available, how to pre-test questionnaires, how to analyze questionnaire results, and other features."
570AD (11 transparencies, teacher guide, sample questionnaires) .. $99.95

MEASURES OF CENTRAL TENDENCY
"Covers the concepts of mean, median, and mode with their methods of calculation, characteristics, and proper use."
W302TAD (10 transparencies, workbook, teacher guide) ............. $99.95

MEASURES OF DISPERSION
"Deals with concepts of range, mean deviation, variance, and standard deviation with other methods of calculation, characteristics, and use."
W303TAD (13 transparencies, workbook, teacher guide) ............. $99.95

THE NORMAL DISTRIBUTION
"One of the most important distributions in statistics is discussed and applications are presented."
W216TAD (10 transparencies, workbook, teacher guide) ............ $109.95

APPLICATIONS OF RANDOM SAMPLINGS
"Covers polls, TV ratings, work sampling, sampling applications in quality control, and other areas. Presents uses of sampling by U.S. Bureau of the Census and other government agencies."
W341TAD (18 transparencies, booklet, teacher guide) ............. $149.95

REGRESSION AND CORRELATION ANALYSIS

"Teaches the principles and applications of these important statistical tools. No knowledge of statistics is required to understand the material. The student is first shown how to plot a scatter diagram and handfit a regression line. After an illustration of the least squares criterion, goodness of fit is illustrated. Regression coefficients are estimated and calculated. Explanations and illustrations of standard error of the estimate, curvilinear regression, coefficients of determination, and multiple regression and correlation are shown."

W102TAD (18 transparencies, teacher guide) ...................... $175.95

ANALYZING TIME SERIES

"Covers essentials of Time Series Analysis, including defining and recognizing time series, why time series analysis is important, components of a time series, recognizing linear and non-linear trends, expressing linear trends mathematically, establishing trendlines through handfitting and the semi-average method, calculating least-squares trend lines, moving averages and their characteristics, seasonal indexes, deseasonalizing time series, calculating seasonal indexes, determining cyclical and irregular components, and forecasting through the use of time series."

1570AD (19 transparencies, teacher guide) ...................... $179.95

ABUSES OF STATISTICS

"Transparencies highlight many of the traps into which uninformed students might fall if they are not aware of the ways in which statistics can be misused, including creating a misleading impression, looking at the wrong thing, not telling the whole story, misusing the average, and inferring cause and effect relationships, among others."

W218TAD (20 transparencies, teacher guide) ...................... $179.95

COMPLETE COURSE IN STATISTICS

"Designed to present the material normally covered in an introductory statistics course, this material requires only a minimum level of mathematical sophistication. Focus is on statistical concepts and their applications. Handouts for the students are included on each topic. Due to the volume of materials in this kit, it will not be available on a preview basis, but a previewing packet is available on request." Includes: Effective Presentation of Data (W215TAD), Frequency Distributions (W301TAD), Measures of Central Tendency (W302TAD), Measures of Dispersion (W303TAD), Probability (W107TAD), The Normal Distribution (W216TAD), Random Sampling (W340TAD), Applications of Random Sampling (W341TAD), Hypothesis Testing (1429AD), Regression and Correlation Analysis (W102TAD), Chi-Square Test (WS301TAD), Analysis of Variance (W306TAD), and Abuses of Statistics (W218TAD).

W406KAD (195 transparencies, random number dice, binomial demonstrators, teacher guide, workbooks, reference materials) .. $1,824.00

3M Company:

| | |
|---|---|
| Introduction to Probability | 15-0245-9 |
| Analytic Geometry - The Straight Line | 15-0253-3 |
| Analyzing Graphs and Charts | 15-2563-3 |

(Each title consists of a packet of 20 originals from which to make transparencies. $2.75/title)

United Transparencies, Inc.:

| | | |
|---|---|---|
| Basic Statistics (23 transparencies with teacher guide) | BS-S | $36.30 |
| Tabulating Test Scores in a Frequency Distribution | Ma 2400 | 1.65 |
| Tabulating Test Scores in a Short Method, 10 pt., c.i. | Ma 2401 | 1.65 |
| Tabulating Test Scores Short Method, 5 pt., c.i. | Ma 2402 | 1.65 |
| The Frequency Polygon | Ma 2403 | 1.65 |
| The Histogram | Ma 2404 | 1.65 |
| Cumulative Frequency | Ma 2[illegible] | 1.65 |
| Cumulative Frequency Curve | Ma 2406 | 1.65 |
| Theoretical Normal Curve | Ma 2407 | 1.65 |
| Negative Skewness | Ma 2408 | 1.65 |
| Positive Skewness | Ma 2409 | 1.65 |
| Frequency Distribution | Ma 2410 | 1.65 |
| Scores in the Interval | Ma 2411 | 1.65 |
| Arithmetic Mean | Ma 2412 | 1.65 |
| Statistical Mean (Language Scores) | Ma 2413 | 1.65 |
| Statistical Mean (Reading Scores) | Ma 2414 | 1.65 |
| Median Scores | Ma 2415 | 1.65 |
| Percentile | Ma 2416 | 1.65 |
| Standard Deviation | Ma 2417 | 1.65 |
| Correlation Coefficient | Ma 2418 | 1.65 |
| Normalized T. Score | Ma 2419 | 1.65 |
| Relation Between Various Types of Derived Scores | Ma 2420 | 1.65 |
| Typical Standard Scores | Ma 2421 | 1.65 |
| Complete Normal Curve | Ma 2422 | 1.65 |

Lansford Publishing Co., Inc.

WALLCHARTS FOR STATISTICS (2060AD (8 wallcharts sold only as set).$89.95)
Each chart is 35 by 45 inches. Printed on heavy stock.

1. The effect of sample size and shape of universe on the distribution of means of random samples.
2. Table of areas under the normal curve.
3. The normal distribution.
4. The use of the table of t-values.
5. Effect of sample size on the distribution of means.
6. The significance of the coefficient of correlation.
7. Validity of psychological tests.
8. Semi-logarithmic charts and arithmetic charts.

LaPine Scientific Company

PROBABILITY AND STATISTICS LAB (Z-08864E ... $32.00)
Materials provide an introduction to demonstration of elements of probability and statistics by a series of 27 suggested experiments described in 120-page manual. Features coin tossing machine which can be set to toss heads continuously, or with adjustment can be altered to 50% probability or other intermediate value. Also included are sampling boxes, a set of dice, numbered cards, a Hexstat unit, an arrow spinner, a distance scale in inches and tenths, and other material correlated to text. Chapter titles include: Even Scattering; Uneven Scattering; Random Numbers; Conclusions from Samples; The Normal Distribution; Random Digits.

HEXSTAT PROBABILITY AND STATISTICS UNIT (Z-08863E .... $6.86)
When Hexstat is held upright, 300 small steel balls pour down [illegible]ugh a network of hexagonal channels and are distributed into 9 paral[illegible] columns to varying heights, producing a bellshaped distribution curve. Repeated readings verify the results. With teacher's manual.

NORMAL DISTRIBUTION CURVE (Z-03964E ... $23.00)
For fast graphing and accurate chalkboard drawing of the curve illustrating the normal probability distribution. Base of curve, 80 cm; height, 40 cm.

NORMAL CURVE TEMPLATE (Z-03969E .... $2.00)
Size 6" x 3½" x 0.030".

VARIATION KIT: NORMAL DISTRIBUTION OF
CHARACTERISTICS OF A SPECIES (Z-08834E ... $15.50)
Introduces the fundamentals of statistics and "real" data. An attempt is made to determine whether plant populations exhibit the same general kind of variations as is found in humans. Contains: 30 packets of sunflower seeds; 22 uniformity containers; 22 container channels; 1 vinyl mount; 1 plastic storage bag; 1 pad of data charts, inst[illegible]ction manual. 30-student kit.

PUBLISHERS

Addison-Wesley Publishing Co.
Jacob Way
Reading, Massachusetts 01867

American Statistical Association
Suite 640, 806 Fifteenth Street
Washington, D.C. 20005

Benjamin/Cummings Publishing Co.
2725 Sand Hill Road
Menlo Park, California 94025

Blackie and Son, Ltd.
450/452 Edgware Road
London W2 1EG
UNITED KINGDOM

Conference Board of the Mathematical Sciences
832 Joseph Henry Building
2100 Pennsylvania Avenue, N.W.
Washington, D.C. 20037

Ginn and Company
A Xerox Corporation
191 Spring Street
Lexington, Massachusetts 02173

Holden-Day, Inc.
San Francisco, California

Holt, Rinehart & Winston
383 Madison Avenue
New York, New York 10017

John Wiley and Sons, Inc.
605 Third Avenue
New York, New York 10016

J. Weston Walch, Publisher
321 Valley Street
Portland, Maine 04104

McMillan Publishing Co., Inc.
866 Third Avenue
New York, New York 10022

McGraw-Hill Book Co.
1221 Avenue of the Americas
New York, New York 10020

National Council of Teachers of Mathematics
1906 Association Drive
Reston, Virginia 22091

Oxford University Press
200 Madison Avenue
New York, New York 10016

Prentice-Hall, Inc.
Englewood Cliffs, New Jersey 07632

Schools Council Project on Statistical Education
Department of Probability and Statistics
The University
Sheffiend S10 2TN
UNITED KINGDOM

Scott, Foresman and Company
1900 East Lake Avenue
Glenview, Illinois 60025

Silver Burdett Company
250 James Street
Morristown, New Jersey 07960

W. H. Freeman and Co., Publishers
669 Market Street
San Francisco, California 94104

W. & R. Chambers, Ltd.
11 Thistle Street
Edinburgh 2
SCOTLAND

W. W. Norton and Company, Inc.
500 Fifth Avenue
New York, New York 10031

Wadworth Publishing Co., Inc.
10 Davis Drive
Belmont, California 94002

# SOURCES OF AUDIO VISUAL MATERIALS

| | Company | Preview Policy |
|---|---|---|
| 1. | Film Services of Hawaii<br>205 Kalihi Street<br>Honolulu, Hawaii 96813<br>(808) 848-0809 | All films are available to educators on a free loan service. |
| 2. | Films, Incorporated<br>1144 Wilmette Avenue<br>Wilmette, Illinois 60091 | No previews are available for rental consideration. Audiovisual directors may preview on 16 mm for purchase determination. |
| 3. | The Food Processors Institute<br>1133 20th Street NW<br>Washington, D.C. 20036<br>(202) 857-0890 | |
| 4. | GAF Corporation<br>Audio/Visual Order Department<br>Binghamton, New York 13902 | All items are shipped on 30-day approval, but school will be invoiced during the approval period. Materials may be returned within the approval period. |
| 5. | Great Plains National Instructional Television Library<br>Box 80669<br>Lincoln, Nebraska 68501<br>(402) 472-2007 | Preview kits available. |
| 6. | Indiana University<br>Audio-Visual Center<br>Bloomington, Indiana 47401<br>(812) 337-2105 | No previews--rental rate covers transportation charges only. |
| 7. | Lansford Publishing Company<br>1088 Lincoln Avenue<br>P.O. Box 8711<br>San Jose, California 95155 | Requests for previews will be charged a $10 per item restocking fee. |
| 8. | LaPine Scientific Company<br>Department E4<br>920 Parker Street<br>Berkeley, California 94710<br>(415) 845-3614 | |
| 9. | Michigan State University<br>Instructional Media Center<br>East Lansing, Michigan 48824<br>(517) 353-4473 | |

| | Company | Preview Policy |
|---|---|---|
| 10. | Prismatron Productions, Inc.<br>155 Buena Vista Avenue<br>Mill Valley, California 94941 | All programs are available on approval free of charge or obligation for a period of 30 days. |
| 11. | 3M Company<br>Visual Products Division<br>3M Center<br>St. Paul, Minnesota 55101 | |
| 12. | United Transparency, Inc.<br>P.O. Box 688<br>Binghamton, New York 13902 | Preview filmstrips are available for 15 days on most sets of transparencies. There is no charge. |

"END OF DOCUMENT"