

From **Brunelli, Lina & Cicchitelli, Giuseppe (editors). Proceedings of the First Scientific Meeting (of the IASE).** Università di Perugia (Italy), 1994. Pages 63-69. Copyright holder: University of Perugia. Permission granted by Dipartimento di Scienze Statistiche to the IASE to make this book freely available on the Internet. This pdf file is from the IASE website at <http://www.stat.auckland.nz/~iase/publications/proc1993>. Copies of the complete Proceedings are available for 10 Euros from the ISI (International Statistical Institute). See <http://isi.cbs.nl/sale-iase.htm> for details.

63

STATISTICS: A MOTIVATOR FOR SECONDARY SCHOOL MATHEMATICS

Richard L. Scheaffer
Department of Statistics, University of Florida
Gainesville, Florida 32611, USA

1. Background

The notion that data collection and analysis should be an integral part of the secondary mathematics curriculum in the United States has enjoyed an ongoing debate since at least the early part of this century. It was suggested by some around the time that Walter Shewhart introduced the control chart for quality improvement in the 1920's and was still being suggested by many in the years following World War II. At that time, one of those proposing that data analysis skills are essential requirements for all students was W. Edwards Deming who, of course, went on to become famous for his work in process improvement. This practical side of mathematics education lost momentum in the 1960's with the advent of the so-called "new math" movement, which emphasized formal mathematical structure and logic.

Not giving up easily, those who thought that students should be exposed to real, everyday applications of mathematics continued to press for more statistical concepts and techniques to be included in the school mathematics curriculum. One of the first organized efforts in this regard came about in the late 1960's with the formation of the Joint Committee on the Curriculum in Statistics and Probability by the American Statistical Association (ASA) and the National Council of Teachers of Mathematics (NCTM). Through publications of materials for a general audience, such as high school students and teachers, and workshops for teachers, this Committee kept honing the conceptual setting and the pedagogical style that would make statistics an acceptable, and even welcome, addition to the mathematics curriculum. Statistics as an educational strand, kept alive by the work of the Joint Committee and others through the 1970's, blossomed into new and vibrant life in the 1980's. This new life for statistics in education parallels the strong interest in process improvement techniques developing in industry during this time, and this duality between the educational and industrial roles for statistics seems to be more than coincidence.

A synergistic mixture of educational policy recommendations, new emphases in statistical practice, and the computer laid new groundwork for statistics in *K-12* education over the last decade. With regard to policy, the major shift in ideas on mathematics education is, perhaps, best summarized in the report *Everybody Counts*, issued by the National Research Council in 1989. This report points out that a transition must be made from preoccupation with routine skills to "developing broad-based mathematical power", as in "using tables, graphs, spreadsheets, and statistical techniques to organize, interpret and present numerical information". Simultaneously, NCTM issued its *Curriculum and Evaluation Standards for School Mathematics*, which contain a carefully delineated strand in statistics throughout the curriculum, as well as an emphasis upon modeling from data in the study of algebra and functions. The data analysis tone of the Standards can be seen in the following quote.

In this age of information and technology, an ever-increasing need exists to understand how information is processed and translated into usable knowledge. Because of society's expanding use of data for prediction and decision making, it is important that students develop an understanding of the concepts and processes used in data analysis.

This standard should not be viewed as advocating, or even prescribing, a statistics course; rather, it describes topics that should be integrated with other mathematics topics and disciplines.

These recommendations for adding statistics to the curriculum are best implemented by having students actively involved with data collection and analysis, and this approach links directly to the emphasis upon data exploration in the practice of modern statistics. The tools of exploratory data analysis are largely graphical, thus requiring the use of a computer for efficient implementation. For the first time, educational policy, the practice of statistics by professionals, and the computer revolution seem to be working in harmony with one another. Data analysis appears to be far more than a passing fad in mathematics education.

Moving to more specific observations, it is true that those writing the statistics strand for the NCTM *Standards* developed many of their ideas around the work being done at that time by the ASA-NCTM Joint Committee. This work was mainly through the Quantitative Literacy Project (QLP), a project funded by the National Science Foundation (NSF) to develop curriculum materials and provide workshops for teachers in the area of statistics for the secondary mathematics curriculum. The Data-Driven Curriculum Project (DDC), which is the focus of the remainder of this report, can be thought of as an extension of the QLP in the directions recommended by the *Standards*.

2. Motivation and need

The technological society of today is dependent upon the collection and interpretation of data for virtually all of its decision making processes. Sample surveys measure consumer preferences and the unemployment rate. Clinical trials determine which new drugs are allowed on the market. Statistical process control improves manufacturing processes and helps keep an industry viable in a competitive world economy. Everyone, then, must understand something about data collection and interpretation if they are to make informed decisions on economics, jobs, politics, health and even entertainment. It is vitally important that the students of today be taught data analysis skills before they become the workers and leaders of tomorrow.

In addition to being important in their own right, data analysis skills help the students build connections between mathematics and other subjects in the school curriculum and to the world outside of the classroom. Data on genetics can lead to interest in the biological sciences. Data on the motion of the stars and planets can lead to interest in the physical sciences. Data on demographics and voting patterns can enhance interest in the social sciences. Data on industrial productivity can strengthen interest in business and economics. These connections can change student attitudes about mathematics and its usefulness as a tool in practical problem solving. Mathematics then becomes both relevant and interesting. Along the way, data collection and summarization can hone basic mathematical skills while data interpretation and model building are developing higher order thinking skills.

Historically, integrating data analysis into the mathematics curriculum has not been a common practice and has not been part of the educational background of mathematics teachers. The materials and inservice workshops of the QLP and similar projects that have sprung up over the past ten years have been well received by teachers, many of whom have attempted to teach the material to their students. Virtually all of these teachers testify to the fact that it takes great effort on their part to fit statistical ideas into the taught curriculum, and some never do succeed. Clearly, there is a need for curriculum materials that embed the important statistical ideas into the traditional mathematics topics so that students see the benefit of both.

3. Goals and objectives

Motivated by the success of the QLP and mindful of the continuing

need for curriculum materials and teacher enhancement in the area of statistics, ASA and NCTM joined forces once more to plan and carry out the Data-Driven Curriculum Project (DDC). The QLP provides materials in certain areas of data exploration, probability, and inference for proportions arising from sample surveys in a style that makes the material accessible to teachers and students at the secondary level. It is not, however, comprehensive in its coverage of statistical topics and it does not allow for easy implementation into the traditional mathematics curriculum. The DDC expands the coverage of statistical topics so that a more comprehensive view of the field emerges and introduces the statistical ideas as motivation for the mathematical topics to which they are tied most closely. Most of the mathematical topics emphasized in the Standards (mainly algebra, functions and geometry) can be illustrated and motivated in this way. Through this two-pronged approach, teachers will be able to use data analysis to meet two objectives: teach a comprehensive body of knowledge in elementary statistics and motivate the major topics in mathematics through real applications.

The specific goals of the DDC are to:

- a) Deliver a detailed scope and sequence plan for a data analysis strand in the high school (9-12) mathematics curriculum that implements the NCTM Standards by developing practical connections to the main concepts of algebra, geometry and functions.
- b) Develop a series of curricular modules fitting the scope and sequence plan in a manner that ensures the coverage of a comprehensive body of knowledge in statistics as well as motivation for many topics of the mathematics curriculum.
- c) Identify materials, including books, computer software, videotapes and manipulatives, that can be used to support the data analysis strand.
- d) Produce a series of workshops for state mathematics supervisors, district coordinators and lead teachers for the purpose of disseminating the results of this project in a manner that will lead to implementation in the schools.

To date, a series of nine modules has been outlined and working drafts have been written for seven of these. Field tests of three modules have been completed and the next four will be field tested in the fall of 1993. A series of five workshops will begin in the fall of 1993. All work is to be completed by the end of 1994.

4. Overview of materials

A brief overview of each of the modules under development, along with the current working titles, is provided below. Much of the work is presented in the form of activities for students working in cooperative groups, with the instructor acting as a catalyst and coach, whereby the students can "discover" key mathematical and statistical ideas on their own.

Our numerical world introduces students to the notion of looking at numbers in the context of what they represent (data) and always thinking about how to interpret these numbers. The unit is built around developing the concepts of size, unit, variability and bias. Activities range from collecting and interpreting data on roller skating injuries to deciding which compact disk player to purchase.

Symbolic expression is built on the fact that a key component of mathematical reasoning is the ability to think abstractly about quantitative concepts so that ideas can be generalized beyond the numbers at hand. This abstraction depends upon symbols and their use in expressions, equations, and functions. A "spreadsheet" formulation of problems is used so that symbols acquire a natural meaning as column labels. Symbolic expressions then show how one column is related to another, independent of the data points in any particular row. Examples include work with wage and hour log sheets for employees in a small firm, discussions involving rates commonly found in vital statistics (birth rates, death rates, accident rates), and development of standard statistical measures of center and spread.

Linear relations builds on the notion that many relationships between continuous variables are studied in reference to a linear function. Data that exhibit approximate linear association are used to illustrate slope as a rate of change. Two such relationships lead to the problem of solving systems of linear equations. Applications relate to changes in automobile prices over time, the association between heights of children and heights of their parents, the relationship between points in a basketball game and percentage of shots made, and the relationship between fat and caloric content of fast foods, to name a few.

Exploring equations extends the concepts begun in the preceding module to develop the idea of an optimal (or "best") line to express the linear relationship between two variables, based on the data. Quadratic functions

and their properties are introduced in the context of the least squares fitting procedure, where optimal properties are illustrated without reference to calculus. Correlation is also discussed.

Mathematical modeling with functions brings together the study of various kinds of functions and the idea of building a mathematical model suggested by data. Students are encouraged to make conjectures about models and then verify or refute those conjectures. Geometric relationships, direct and inverse variation, exponential and logarithmic functions, and proportionality are the mathematical topics of emphasis. Fitting models to data, transforming data, and studying patterns in residuals are some of the statistical ideas considered. Examples include exponential growth and decay of certain populations, relationships between weight and size of animals, and growth of plants over time.

Centers considers the notion of “average” from a geometric perspective, noting that most definitions of “center” actually minimize some distance measure. Weighted averages and the centroids of planar figures are given meaning through realistic applications such as determining a basketball player's performance rating and locating a warehouse to minimize travel distance to stores.

Frequency tables and probability shows that data in the form of relative frequencies for categorical responses can be interpreted as approximate probabilities for predicting future events of a similar nature. Thus, marginal, joint and conditional relative frequencies are used to motivate similar concepts for probability. Testing hypotheses (chi-square test) is introduced here through simulation arguments, and simulation is used to introduce a variety of probability distributions for discrete random variables. A number of examples are drawn from clinical trials experiments and drug screening tests.

Advanced modeling develops matrix algebra through extending many of the modeling notions from earlier modules to more complex cases. Statistical models that go beyond simple linear regression are presented, as are probabilistic models involving Markov chains.

Projects outlines the essential steps that cooperative groups should consider when attempting a project that goes from posing a question, through data collection and analysis, and on to a written interpretation of results. Guidelines are given for designing sample surveys and experiments, with the latter concentrating on “clinical trials” and factorial arrangements of

treatments. Students are encouraged to critique articles from the media that include statistical arguments.

Taken together, the nine modules cover statistical topics on exploring data, collecting data through planned studies, and inference. These are spread out over modules that could be used from grade 9 through grade 12 in covering a majority of the mathematical topics deemed worthy of enhanced emphasis by NCTM. It is hoped that this project will successfully demonstrate that statistics and mathematics can be integrated into a modern curriculum and will serve as a model for future curriculum development.

5. Summary

The development of a data-driven curriculum for high school mathematics appears to be in line with the needs of students to see more motivation and application within the mathematics classroom and to develop important skills to carry beyond the classroom. The revised curriculum under development is designed to raise the quantitative literacy of all students as it builds connections among mathematics, science, and technology. It models a new approach to the teaching of mathematics, the approach required by the NCTM *Standards*, as it emphasizes hands-on activities for students and discovery of concepts through data. Technology in the form of graphing calculators and computers is an integral part of the teaching and learning style being promoted through these materials and workshops. This project attempts to connect topics of importance in a modern mathematics curriculum to a modern view of statistical science for the purpose of enhancing student interest and skills in both areas.