NCTM Standards in Action
The Content Standard: Data Analysis and Probability

Klaus Puhlmann

The question, "What mathematical content and processes should students know, understand and be able to use in the K–12 curriculum?" is an ongoing preoccupation with curriculum developers. The National Council of Teachers of Mathematic's *Principles and Standards for School Mathematics* (NCTM 2000) strongly recommends 10 content standards for the K–12 school mathematics curriculum, one of which is data analysis and probability. Thus, data analysis and probability as a content strand is included at every level, albeit with varying degrees of depth and breadth. It is well articulated across the grades, thus ensuring the accumulation of important ideas and building successively deeper and more refined understanding.

The data analysis and probability standard identifies the broad areas of emphasis that enable all K–12 students to

- formulate questions that can be addressed with data and collect, organize and display relevant data to answer them;
- select and use appropriate statistical methods to analyze data;
- develop and evaluate inferences and predictions that are based on data; and
- understand and apply basic concepts of probability.

There is no doubt that our lives are inundated with large quantities of data. Data are often the basis for decision making in business, industry, politics, research and even in our personal lives. Consumer surveys guide marketing strategies; purchasing products, which is often based on data related to satisfaction levels about a particular product; and decisions that are often based on probabilistic reasoning. Therefore, it is important that students know about data analysis and probability to reason statistically and become informed citizens and intelligent consumers.

There are, of course, other valid reasons for including data analysis and probability as a content strand across the grades. The data analysis and probability strand allows teachers and students to make several important connections among ideas and procedures, from number, algebra, measurement and geometry. It also offers natural ways for students to connect mathematics with other subject areas and events occurring in their own lives. Engaging students in statistical reasoning about data not only serves them well as they enter the world of work and living independently but also teaches them that solutions to some problems depend on assumptions and have some degree of uncertainty. This kind of reasoning used in statistics and probability is not always intuitive; hence, students would not necessarily develop this important skill from other parts of the curriculum.

The expectations of students in this content strand are naturally age-appropriate. At the primary level, the expectations are that students

- pose questions and gather data about themselves and surroundings;
- sort and classify objects according to their attributes and organize data about the objects;
- represent data using concrete objects, pictures and graphs;
- describe parts of the data and the set of data as a whole to determine what the data show; and
- discuss events related to students' experience as likely or unlikely.

At the upper elementary level, students are

- designing investigations and considering how collection methods affect the nature of the data set;

- collecting data using observations, surveys and experiments;
- representing data using tables and graphs (for example, line plots, bar graphs and line graphs);
- describing the shape and important features of a set of data, comparing data and determining how the data are distributed;
- using measures of central tendency;
- comparing different representations of the same data;
- proposing and justifying conclusions and predictions, and designing studies;
- describing events as likely or unlikely, including such words as certain, equally likely and impossible;
- predicting the probability of outcomes; and
- gaining and understanding the likelihood of an event, represented by a number from 0 to 1.

At the middle school level, students are

- formulating questions, designing studies and collecting data about a characteristic shared by two populations or within one population;
- selecting, creating and using appropriate graphical representations of data (for example, histograms, box plots and scatter plots);
- finding, using and interpreting measures of centre and spread;
- discussing and understanding the correspondence between data sets and their graphical representations;
- using observations about differences between two or more samples to make conjectures about the population from which the samples were taken;
- making conjectures about possible relationships between two characteristics of a sample;
- using conjectures to formulate new questions;
- understanding and using appropriate terminology to describe complementary and mutually exclusive events;
- using proportionality and a basic understanding of probability to make and test conjectures; and
- comparing probabilities for simple compound events, using such methods as organized lists, tree diagrams and area models.

At the high school level, students are

- understanding the differences among various kinds of studies and the types of inferences that can be drawn from them;
- knowing the characteristics of well-designed studies, including randomization in surveys and experiments;
- understanding the meaning of measurement data and categorized data, of univariate and bivariate data and of the term *variable*;
- understanding histograms, parallel box plots and scatter plots, and their appropriate use;
- computing basic statistics and understanding the distinction between a statistic and a parameter;
- able to display the distribution for univariate measurement data;
- able to display a scatter plot for bivariate measurement data and determine regression;
- computing coefficients, regression equations and correlation coefficients using technology;
- displaying and discussing bivariate data with one categorized variable;
- recognizing how linear transformations of univariate data affect shape, centre and spread;
- identifying trends in bivariate data and finding functions that model the data;
- using simulations to explore the variability of sample statistics from a known population and constructing sampling distributions;
- understanding how sample statistics reflect the values of population parameters and using sampling distribution as the basis for informal inference;
- evaluating published reports that are based on data with a focus on the design of the study, the appropriateness of the data analysis and the validity of conclusions;
- understanding how basic statistical techniques are used in the workplace;
- understanding the concepts of sample space and probability distribution;
- using simulations to construct empirical probability distributions;
- compiling and interpreting the expected value of random variables;
- understanding the concept of conditional probability and independent events; and
- understanding how to compute the probability of a compound event.

This content standard is well articulated across the grades, moving from the simple notions and concepts to a relatively high level of sophistication at the high-school level. The study of data analysis and probability is based on the premise that students need to have hands-on experiences working with data derived from our everyday lives.

The content standard (data analysis and probability) recognizes the importance of having all students develop an awareness of the concepts and processes of data analysis and probability. It is essential that students learn that data analysis and probability is more than reading and interpreting graphs, but that it is also an effective tool for solving problems. The study of data analysis and probability highlights the importance of questioning, conjecturing and searching for relationships when formulating and solving real-world problems.

Students at all grade levels begin to understand that data come in various forms and in different quantities, requiring different treatment and organization to extrapolate meaningful information. Because of the prevalence of statistical data, summaries, graphs and probabilistic problems in our everyday lives and at work, it is important that data analysis and probability be given a more prominent place in the K–12 mathematics curriculum.

The three articles that follow relate to the data analysis and probability content standard. The first article presents an example from everyday life—applying basic probability concepts in a practical setting. The second article discusses how teachers can teach the concept of a mean in a meaningful way. The third article presents an elementary mathematics research model that allows the students to begin early to collect, organize and describe data. It is founded on problem-finding and problem-solving behaviours and is designed to support the development of high-level thinking as students' thoughts diverge and converge throughout the research process.

Bibliography

Alberta Learning. *Alberta Program of Studies for Grades K–6 Mathematics*. Edmonton: Author, 1996a.

——. *Alberta Program of Studies for Grades 7–9 Mathematics*. Edmonton: Author, 1996b.

——. *Alberta Program of Studies for Grades 10–12 Mathematics*. Edmonton: Author, 2001.

National Council of Teachers of Mathematics (NCTM). *Curriculum and Evaluation Standards for Teaching School Mathematics*. Reston, Va.: Author, 1989.

——. *Principles and Standards for School Mathematics*. Reston, Va.: Author, 2000.

Abul-Wefa (940–998)

The Persian mathematician credited for his improvements in trigonometry posed the following problem:

Two of three congruent squares are to be cut into eight pieces so that, together with the third square, they can be arranged into one larger square.