NCTM Standards in Action

Klaus Puhlmann

The Standards recognize the importance of having all students develop an awareness of the concepts and processes of data analysis, statistics and probability. These concepts are more than just reading and interpreting graphs. Statistics and probability are important links to other content areas, such as social studies and science. They also can reinforce communication skills as the students discuss and write about their activities and their conclusions. Also within mathematics, these topics often involve the uses of number, measurement, estimation and problem solving.

The Standards for K–12 suggest that mathematics instructional programs should include attention to data analysis, statistics and probability so that all students

- pose questions and collect, organize, and represent data to answer these questions;
- interpret data using methods of exploratory data analysis;
- develop and evaluate inferences, predictions, and arguments that are based on data;
- understand and apply basic notions of chance and probability. (NCTM 1998)

Students in the early grades should be exposed to real objects that embody the characteristics to be studied. Activities such as comparing, sorting and counting are essential in order to develop the students' understanding of data and data analysis. Although the depth and variety of activities in relation to the above-mentioned standards differ greatly, the spirit of investigation and exploration is to permeate statistics instruction at all levels. Students' questions about the world around them can often be answered by collecting and analyzing data. As the students develop questions, they also decide what information to collect that will answer these questions. This process also involves evaluating the data collected, interpreting them and drawing conclusions from them. The focus areas for the primary grades include

- gathering data about themselves and their surroundings;
- sorting and classifying objects and organizing data according to attributes;
- representing data to convey results using concrete objects, pictures and numbers;
- describing parts of the data and the data as a whole;
- identifying parts of the data with special characteristics; and
- understanding notions such as *certain, impossible, more likely, less likely*.

Students need to understand and develop an appreciation that the overall purpose of collecting, organizing and representing data is to answer questions that are otherwise difficult to answer. Our teaching needs to build on the informal experiences that the primary students bring to this task. The students must eventually see that when the data are organized and represented, they often convey powerful information that either leads to a conclusion or more questions. This is also an excellent opportunity to engage students in discussions, allowing them to communicate their thoughts and understanding.

Making inferences and predictions is generally deferred to the upper elementary grades, as both require probabilistic thinking. However, answering questions about notions like *certain, impossible, more likely* and *less likely* begins at this level. Again, the treatment of these notions must be built on the intuitive understanding that these primary students bring to these discussions. Therefore, the theoretical treatment of the notion of probability is not appropriate at this grade level.

At the upper-elementary level, students continue to focus on the same standards, but the focus areas increase in variety and depth. Students at this level should

- formulate questions they want to investigate;
- design data investigations to address a question;
- collect data using observations, measurement, surveys, or experiments;
- organize data using tables and graphs (e.g., bar graph, line plot, stem-and-leaf plot, circle graph, and line graph);

- use graphs to analyze data and to present information to an audience;
- compare data representations to determine which aspects of the data they highlight or obscure;
- describe the shape and important features of a set of numerical data, including its range, where the data are concentrated or sparse, and whether there are outliers;
- describe the center of sets of numerical data, first informally, then using the median;
- classify and describe categorical data (e.g., ways we travel to school) in different ways; analyze and compare the information highlighted by different classifications;
- compare related data sets, with emphasis on the range, center, and how the data are distributed;
- propose and justify conclusions based on data;
- formulate questions or hypotheses based on initial data collection, and design further studies to explore them;
- describe how data collection methods can impact the nature of the data set;
- discuss the concept of representativeness of a sample within the context of a particular example (e.g., is the class representative of other fifth-grader classes in our town? In Alberta? In Canada? Why or why not?);
- compare the data from one sample to other samples and consider why there is variability;
- in simple experiments, infer the structure of the population through drawing repeated samples (with replacement);
- discuss events as likely or unlikely and give descriptions of the degree of likelihood in informal terms (e.g., unlikely, very unlikely, certain, impossible);
- estimate, describe, and test probabilities of outcomes by associating the degree of certainty with a value ranging from 0 to 1 (e.g., in simple experiments involving spinners with different fractions shaded). (NCTM 1998)

Students at this grade level must be engaged in activities that involve responding to questions about a variety of situations that interest them. They should also become familiar with a variety of representations and their appropriateness for different data and purposes. The focus at this level must also include looking at data in more than one way, comparing data and developing the idea of "typical" or average value. Furthermore, their experiences must also include comparing several, related data sets and the use of evidence as validation.

Toward the end of this grade level, students should also develop an understanding that data sets are samples of larger populations. At this point, too, they are beginning to experiment with probability in situations with a few outcomes, and they develop the language to discuss their informal notions of probability.

Students at the middle grade level continue to build on their previous experiences, but they are now engaging in the full process of data investigation: posing questions, collecting data, organizing data, analyzing data, interpreting data and answering questions. The students now draw on their knowledge of ratios, fractions, decimals, percent, graphs and measurement as they engage in data analysis.

Furthermore, they develop and extend their understanding of ideas that are central to the study of statistics, such as data distribution, central tendency and variance. New representational forms are being added as well (for example, box-and-whisker plots) to allow more complex considerations of data distributions. Students are also engaged in the analysis of scatter plots for related variables, allowing them for the first time to develop linear approximations (line of best fit) for the plots. With respect to the following focus areas the students should

- design experiments and surveys, and consider potential sources of bias in design and data collection;
- recognize types of data (for example, categorical, count, continuous or measurement) and organize collections of data;
- choose, create and use various graphical representations of data (line plots, bar graphs, stem-and-leaf plots, histograms, scatter plots, circle graphs and box-and-whisker plots) appropriately and effectively;
- find, describe and interpret mean, median and mode as measure of the centre of a data set; know which measure is best to use in particular situations; and understand how each does and does not represent the data;
- describe and interpret the spread of a set of data using tools such as range, interquarterly range and box-and-whiskers graphs;
- interpret graphical representations of data, including description and discussion of the meaning of the shape and features of the graph, such as symmetry, skewness and outliers;
- analyze associations between variables by comparing the centres, spreads and graphical representations of related data sets;
- examine and interpret relationships between two variables using tools such as scatter plots and approximate lines of best fit;
- develop conclusions about a characteristic in the population from a well-constructed sample;
- through simulations, develop an understanding about when differences in data may indicate an

actual difference in the populations from which the data were collected and when the differences may result from natural variation in samples;
- use data to answer the questions that were posed, understand the limitations of those answers and pose new questions that arise from the data;
- make judgments about the likelihood of uncertain events and be able to connect those judgments to percents of proportions;
- understand what it means for events to be equally likely and for a game or process to be fair;
- compute simple probabilities using appropriate methods, such as lists, tree diagrams or area models;
- identify complementary, mutually exclusive, independent and dependent events and understand how these relationships affect the determination of probabilities.

During the final four years of schooling, the complexity and difficulty levels are further enhanced. While the standards continue to be the same, the focus areas are further increasing in depth and breadth. More specifically, students at this level should

- design and carry out appropriate methods for gathering univariate data, both to study the distribution of a variable in one population and to compare the distributions of the same variable in two different populations;
- design appropriate methods for collecting, recording and organizing data to obtain bivariate data to study the association between two variables;
- select appropriate graphical representations and numerical summaries of data;
- understand how a change in a representation (for example, scales on a scatterplot, categories in a two-way table and bin size of a histogram) affects the information it conveys;
- use calculators and computer applications (for example, spreadsheets, simulation software and statistical software) appropriately to assist in data collection, organization and representation;
- compute, identify and interpret measures of centre and spread (for example, range, variance and standard deviation, and interquartile range);
- describe shapes of one- and two-dimensional data sets;
- look for symmetry and skewness, clusters and gaps, and possible outliers in data and consider their effects on the interpretation of the data;
- recognize how sample size or transformations of data affect shape, centre and spread;
- use a variety of representations of data, including scatterplots, frequency distributions and two-way tables;
- be able to recognize trends in bivariate data, visually and numerically, and use technology to determine how well different models (for example, linear, exponential, and quadratic) fit data, while understanding that a perfect fit is unlikely for empirical data;
- understand the elements involved in finding good models for phenomena;
- apply well-fitting models to predict unobserved outcomes;
- evaluate conclusions based on data;
- use data from samples to estimate population statistics;
- use and interpret the normal and binomial distributions appropriately;
- understand and compute probabilities of independent, disjoint and conditional events;
- understand that some phenomena are random and apply and the law of large numbers to predict long-term behavior;
- use probability distributions to compute probabilities of events.

The emphasis at this level has shifted from univariate data to binomial data. Students will also deal with the concepts of randomness and chance in greater depth. Their increased mathematical knowledge will also allow them to formalize their understanding of the relationships among the graph of a distribution, measures of shape, centre and spread. Linear approximations are now including finding regression curves, and technology is used more extensively to study data, illustrate concepts, perform calculations, create representations and provide data from situations.

Students at this level will have studied probability distributions, binomial and normal distributions, designing surveys and experiments, the influence of sample size on the closeness of the estimate of a population statistic and drawing inferences about a population.

The three articles that follow provide two examples of the level of student understanding relative to certain statistical concepts and one example of how a real-life situation can be used to capture students' interest in statistics.

Bibliography

Alberta Education. *The Common Curriculum Framework for K–12 Mathematics*. Edmonton: Author, 1985.

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

——. *Principles and Standards for School Mathematics: Discussion Draft. NCTM Standards 2000.* Reston, Va.: Author, 1998.