Teaching Statistics in K-12 classrooms: Emphasis on Data Analysis

Madhuri S. Mulekar
Department of Mathematics & Statistics
University of South Alabama
Mobile, AL 36688
mmulekar@jaguar1.usouthal.edu

Abstract

Over last decade, NCTM guidelines have indicated need for an added emphasis on data analysis in K-12 mathematics classes. The College Board Standards for College Success in Mathematics and Statistics (2006) describes mathematical knowledge essential for students' success in college, which includes data analysis as an important part of mathematics courses in middle school and high school at six different levels leading to Advanced Placement (AP) courses. Keeping in mind the needs of Pre-K-12 education, the American Statistical Association developed Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report. Since 1997, the College Board started AP exams in Statistics and the need for teachers trained in statistics including data analysis increased. In this session, some of these guidelines will be discussed and time-permitting a data-analysis activity will be conducted.

Guidelines

- NCTM guidelines on data analysis in K-12 mathematics classes.
- The College Board Standards for College Success in Mathematics and Statistics
- ASA Guidelines for Assessment and Instruction in Statistics Education (GAISE)
- Advanced Placement Exam in Statistics

NCTM Standards

Data Analysis and Probability

Instructional programs from pre-kindergarten through grade 12 should enable all 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.

The “Data Analysis and Probability” standard recommends that students formulate questions that can be answered using data and address what is involved in wisely gathering and using that data. Students should learn how to collect data, organize their own or others’ data, and display the data in graphs and charts that will be useful in answering their questions. This standard also includes learning methods for analyzing data and ways of making inferences and drawing conclusions from data. The basic concepts and applications of probability also are addressed, with an emphasis on the way probability and statistics are related.

College Board Standards

- The College Board Standards for College Success in Mathematics and Statistics (2006)
- As part of its commitment to preparing *all* students for college success, opportunity in the workplace, and effective participation in civic life, the College Board developed the College Board Standards for College Success™ to help states, school districts, and schools provide all students with the rigorous education that will prepare them for this success. This commitment is founded on the belief that all students can meet high expectations for academic performance when they are taught to high standards by qualified teachers.
- Preparing students for college *before* they graduate from high school is critical to students' completing a college degree.

College Board Standards

Endorsed by:

- National Council of Teachers of Mathematics (NCTM)
- Mathematical Association of America (MAA)
- American Mathematical Society (AMS)
- American Statistical Association (ASA)
- American Mathematical Association of Two-Year Colleges (AMATYC)
- Achieve, Inc.

Middle School Math I

STANDARDS

- Nonnegative Rational Numbers and Concepts of Integers [Number and Operations]
- Ratios and Rates [Number and Operations]
- Two-Dimensional Geometry and Measurement [Geometry and Measurement]
- Univariate Data Analysis [Data]
- Experimental and Theoretical Probability [Probability]
- Linear Patterns and Relationships [Algebra]

Middle School Math I

- Math I students work on developing the ability to formulate simple questions that can be answered by collecting and analyzing data. They design straightforward data investigations and collect data on one attribute for each member of a small population. Students develop abilities to describe the distribution of elements in such data sets (extremes, mean, median, mode, range, outliers) as well as to identify whether data are categorical or numerical. Students display collected data in graphs or tables, and they interpret the results of their investigation in the context of the formulated question.
- Math I students' experience with probability focuses on simple experiments viewed from experimental and theoretical standpoints. They develop probability as a ratio naming a number between 0 and 1. In their study of content from algebra, these students examine and graph patterns showing linear relations, develop expressions with whole-number coefficients for linear patterns, and describe such patterns using all forms of representation. Students also solve one- step linear equations and evaluate linear expressions for a given whole number.

Standard MI.4: Univariate Data Analysis

Students formulate and answer questions about a small population by collecting and analyzing univariate data from the population. Students communicate the results of data analyses and critique data analyses reported in the media.

Objective MI.4.1:

The student formulates questions about a small population that can be answered through data collection and analysis, designs related data investigations, and collects data.

Performance Expectations:

MI.4.1.1 Formulates a simple question, and defines a small population on which data can be collected to answer the question.

MI.4.1.2 Identifies an attribute on which to collect data, and decides how to measure the attribute to answer the question formulated.

MI.4.1.3 Determines a data collection process appropriate to the question formulated, and collects data on one attribute associated with each member of the population.

MI.4.1.4 Recognizes and describes the differences between numerical and categorical data.

Standard MI.4: Univariate Data Analysis

Objective MI.4.2:

The student organizes and summarizes categorical and numerical data using summary statistics and a variety of graphical displays.

Performance Expectations:

MI.4.2.1 Constructs appropriate graphical displays (bar graphs, circle graphs, line plots, stem-and-leaf plots, histograms), with and without technology, to describe the distribution of data values.

MI.4.2.2 Describes the shape, center, and spread of the distribution of numerical data; constructs frequency distributions and determines the mode for categorical data.

MI.4.2.3 Computes measures of center (mean, median) and spread (range) for a set of numerical data, with and without technology, and explains the influences of outliers on each measure.

Standard MI.4: Univariate Data Analysis

Objective MI.4.3:

The student interprets results and communicates conclusions regarding a formulated question using appropriate symbols, notation, and terminology. The student critiques simple presentations of data and conclusions found in the media.

Performance Expectations:

MI.4.3.1 Interprets results and communicates conclusions in the context of the formulated question using appropriate symbols, notation, and terminology.

MI.4.3.2 Critiques a simple data presentation, such as those found in the media, including an analysis of the reasonableness of the display(s) used and the conclusions drawn.

Standards developed for

- Middle Grades Mathematics
- Middle School Math I
- Middle School Math II
- Introduction to Algebra I through Precalculus
- Algebra I and Data Analysis
- Geometry
- Algebra II and Data Analysis
- Precalculus and Data Analysis

Advanced Placement Statistics

- Colleges tend to offer 2 levels of introductory statistics courses
  - Calculus based
  - Non-calculus based
- AP statistics is considered “equivalent” to the non-calculus based college level introductory statistics course.
- The first exam was given in May 1997.
- One of the fastest growing AP course offering.

  Enrollment: 8,000 in 1997

  Over 100,000 in 2008

General Course Information: AP Statistics

- Equivalent to a one-semester, introductory, non-calculus based college course.
- Majority of high schools teach the course over an entire school year.
- Pre-requisite is Algebra II.
- Use of technology is emphasized by the use of computers/calculators.

Basic Four Conceptual Themes AP Statistics

- Exploring Data
- Planning a Study
- Anticipating Patterns
- Statistical Inference

GAISE Guidelines

- Guidelines for Assessment and Infrastructure in Statistics Education
- Endorsed by the American Statistical Association – August 2005
- http://www.amstat.org/education/GAISE/
- Goal: Statistical literacy

The Difference between Statistics and Mathematics

"Statistics is a methodological discipline. It exists not for itself, but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data. The need for such a discipline arises from the *omnipresence of variability*." (Moore and Cobb, 1997)

A major objective of statistics education is to help students develop statistical thinking. Statistical thinking, in large part, must deal with this omnipresence of variability; statistical problem solving and decision making depend on understanding, explaining, and quantifying the variability in the data.

It is this focus on *variability* in data that sets apart statistics from mathematics.

The Nature of Variability

Measurement Variability — Repeated measurements on the same individual vary. Sometimes two measurements vary because the measuring device produces unreliable results, such as when we try to measure a large distance with a small ruler. At other times, variability results from changes in the system being measured. For example, even with a precise measuring device, your recorded blood pressure could differ from one moment to the next.

The Nature of Variability

Natural Variability — Variability is inherent in nature. Individuals are different. When we measure the same quantity across several individuals, we are bound to get differences in the measurements. Although some of this may be due to our measuring instrument, most of it is simply due to the fact that individuals differ. People naturally have different heights, different aptitudes and abilities, and different opinions and emotional responses. When we measure any one of these traits, we are bound to get variability in the measurements. Different seeds for the same variety of bean will grow to different sizes when subjected to the same environment because no two seeds are exactly alike; there is bound to be variability from seed to seed in the measurements of growth.

The Nature of Variability

Induced Variability — If we plant one pack of bean seeds in one field, and another pack of seeds in another location with a different climate, then an observed difference in growth among the seeds in one location with those in the other might be due to inherent differences in the seeds (natural variability), or the observed difference might be due to the fact that the locations are not the same. If one type of fertilizer is used on one fi eld and another type on the other, then observed differences might be due to the difference in fertilizers. For that matter, the observed difference might be due to a factor we haven't even thought about. A more carefully designed experiment can help us determine the effects of different factors.

The Nature of Variability

Sampling Variability — In a political poll, it seems reasonable to use the proportion of voters surveyed (a sample statistic) as an estimate of the unknown proportion of all voters who support a particular candidate. But if a second sample of the same size is used, it is almost certain that there would not be exactly the same proportion of voters in the sample who support the candidate. The value of the sample proportion will vary from sample to sample. This is called sampling variability. So what is to keep one sample from estimating that the true proportion is .60 and another from saying it is .40? This is possible, but unlikely, if proper sampling techniques are used. Poll results are useful because these techniques and an adequate sample size can ensure that unacceptable differences among samples are quite unlikely.

Role of context and probability

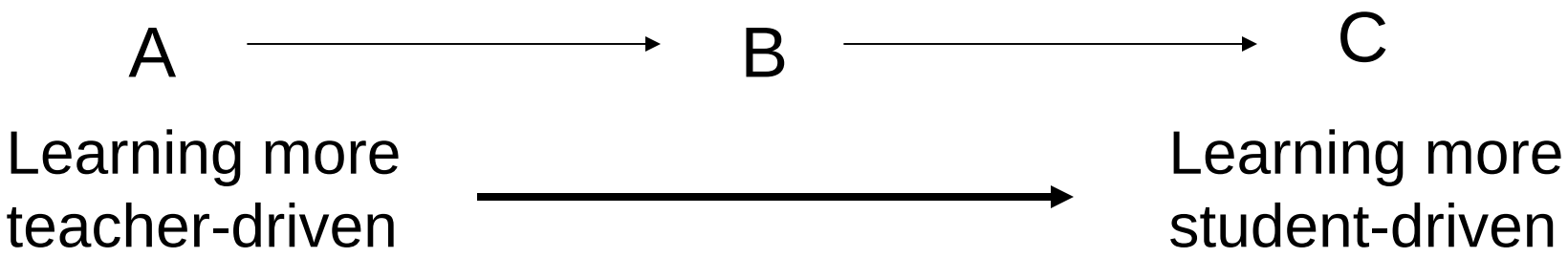
Statistics requires a different kind of thinking, because ***data are not just numbers, they are numbers with a context.*** In mathematics, context obscures structure. In data analysis, context provides meaning."

Probability is a tool for statistics.

Steps of the Problem Solving

- Formulate Questions
- Collect Data
- Analyze data
- Interpret Results

Three different Developmental Levels: A, B, C



- Discussion
- Activity