A FRAMEWORK FOR TEACHING STATISTICS WITHIN THE K-12 MATHEMATICS CURRICULUM

Gary D. Kader and Mike Perry
Appalachian State University, United States
gdk@math.appstate.edu

This paper reports on an American Statistical Association project which developed ASA-endorsed guidelines for teaching and learning statistics at the Pre K-12 level. A group of leading statistics and mathematics educators developed the report, "A Curriculum Framework for Pre K-12 Statistics Education." These guidelines complement the NCTM Principals and Standards of School Mathematics – providing additional guidance and clarity on the data analysis strand. A major goal of the document is to describe a statistically literate high school graduate and, through a connected curriculum, provide steps to achieve this goal. Topics for discussion include:

- *Developing statistical literacy within the Pre K-12 mathematics curriculum*
- *Links to the NCTM Standards*
- *Impact of high stakes testing*
- *Differences between mathematics and statistics*
- *Key components and concepts associated with the data analysis process*
- *Examples illustrating connections in key statistical concepts across all grade levels*

INTRODUCTION

In 2005, the Board of Directors of the American Statistical Association (ASA) endorsed the report *A Curriculum Framework for Pre K-12 Statistics Education* (2005). The development of this *Framework* was supported by the ASA through funding of a Strategic Initiative Grant proposed by the ASA Advisory Committee on Teacher Enhancement. It is designed to give educators guidance towards developing statistically literate citizens. The writers were Christine Franklin, Gary Kader, Denise Mewborn, Jerry Moreno, Mike Perry, Roxy Peck, and Richard Scheaffer.

The main objectives of the *Framework* are to provide a conceptual structure for Pre K-12 statistics education and to provide guidelines toward developing statistically literate citizens. Its foundation rests on the National Council of Teachers of Mathematics' (NCTM) *Principles and Standards for School Mathematics* (2000). *The Framework* presents a coherent model describing a conceptual and developmental structure for the overall statistics curriculum at the Pre K-12 level. It is designed to provide guidance for educators including writers of state standards, writers of assessment items, educators at teacher preparation programs, curriculum directors, and Pre K-12 teachers.

THE ULTIMATE GOAL: STATISTICAL LITERACY

Statistical literacy is essential in our personal lives as consumers, citizens and professionals. Statistics plays a role in our health and happiness. Sound statistical reasoning skills take a long time to develop. They cannot be honed to the level needed in the modern world through one high school course. The surest way to reach the necessary skill level is to begin the educational process in the elementary grades and keep strengthening and developing these skills throughout the middle and high school years. A statistically literate high school graduate will know how to interpret the data in the morning newspaper and will ask the right questions about statistical claims. He or she will be comfortable handling quantitative decisions that occur on the job, and will be able to make informed decisions about the quality of life issues.

NCTM STANDARDS AND THE FRAMEWORK

The *Curriculum and Evaluation Standards for School Mathematics* developed by the National Council of Teachers of Mathematics (NCTM) was published in 1989 and included Data Analysis and Probability as one of its five content strands. This document and its 2000 replacement entitled *Principles and Standards for School Mathematics* became the basis for reform of mathematics curricula in many states of the USA. The NCTM *Principles and Standards* describes the statistics content strand as follows.

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;
- understand and apply basic concepts of probability."

The NCTM document elaborates on these themes somewhat and provides examples of the types of lessons and activities that might be used in a classroom.

Statistics, however, is a relatively new subject for many teachers who have not had an opportunity to develop sound knowledge of the principles and concepts underlying the practices of data analysis that they are now called upon to teach. These teachers do not clearly understand the difference between statistics and mathematics. They do not see the statistics curriculum for grades K-12 as a cohesive and coherent curriculum strand. These teachers may not see how the overall statistics curriculum provides a developmental sequence of learning experiences. The *Framework* provides a conceptual structure for statistics education which gives a coherent picture of the overall curriculum. This structure adds to but does not replace the NCTM recommendations.

THE DIFFERENCE BETWEEN STATISTICS AND MATHEMATICS

A major objective of statistics education is to help students develop statistical thinking. Statistical thinking, in large part, must deal with the omnipresence of variability.

> "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*." (Cobb and Moore, 2000)

Statistical problem solving and decision making depend upon understanding, explaining and quantifying the variability in data. It is this focus on *variability in data* that sets statistics apart from mathematics. There are many different sources of variability in data. Some of the important sources are described below.

- *Measurement Variability*
  Repeated measurements on the same individual may vary. Sometimes two measurements vary because the measuring device produces unreliable results. Other times variability results from changes in the system being measured.
- *Natural Variability*
  Variability is inherent in nature. Individuals are different. When the same quantity is measured across several individuals, there is bound to be some differences in the measurements.
- *Induced Variability*
  The fundamental concept of experimental design is comparison. This one basic idea of comparing natural variability to the variability induced by other factors forms the heart of modern statistics.
- *Sampling Variability*
  Different samples may give different results. This sample to sample variability is the basic issue which must be confronted in the use of surveys.

Many mathematics problems arise from applied contexts, but the context is removed to reveal mathematical patterns.

> "In mathematics, context obscures structure. In data analysis, context provides meaning." (Cobb and Moore, 2000)

Statisticians, like mathematicians, look for patterns, but the meaning of the patterns depends on the context.

> "The focus on variability naturally gives statistics a particular content that sets it apart from mathematics itself and from other mathematical sciences, but there is more than just content that distinguishes statistical thinking from mathematics. Statistics requires a different *kind* of thinking, because *data are not just numbers, they are numbers with a context."* (Cobb and Moore, 2000)

PROBABILITY

Probability as a Tool for Statistics

Probability is an important part of any mathematical education. It is a part of mathematics that enriches the subject as a whole by its interactions with other uses of mathematics. Probability is an essential tool in applied mathematics and mathematical modeling. It is also an essential tool in statistics. But the use of probability as a mathematical model and the use of probability as a tool in statistics employ not only different approaches but different kinds of reasoning. Two problems and the nature of the solutions will illustrate the difference.

> Problem 1
> Assume a coin is "fair."
> Question: If we toss the coin 5 times, how many heads will we get?

> Problem 2
> You pick up a coin.
> Question: Is this a fair coin?

Problem 1 is a mathematical probability problem. Problem 2 is a statistics problem, which can use the mathematical probability model determined in problem 1 as a tool to seek a statistical solution.

Probability and Chance Variability

Two important uses of "randomization" in statistical work occur in sampling and experimental design. When sampling we "select at random" and in experiments we "randomly assign individuals to different treatments." Randomization does much more than remove bias in selections and assignments. Randomization leads to chance variability in outcomes, which can be described with probability models. When randomness is imposed, the statistician wants to know if the observed result is due to chance variation or something else? This is the idea of statistical significance.

THE ROLE OF MATHEMATICS IN STATISTICAL EDUCATION

The evidence that statistics is different from mathematics is not presented to argue that mathematics is not important to statistics education or that statistics education should not be a part of mathematics education. To the contrary, statistics education becomes increasingly mathematical as the level of understanding increases. But data collection design, exploration of data, and the interpretation of results should be emphasized in statistics education for statistical literacy. These are heavily dependent on context, but at the introductory level involve limited formal mathematics. Probability plays an important role in statistical problem solving, but formal mathematical probability should have its own place in the curriculum. Pre-college statistics education should emphasize the ways that probability is used in statistical thinking; an intuitive grasp of probability will usually suffice at these levels.

THE FRAMEWORK

Underlying Principles

The *Framework* presents statistical problem solving as an investigative process that involves four components:

Formulate Questions
- clarify the problem at hand
- formulate one (or more) questions that can be answered with data

Collect Data
- design a plan to collect appropriate data
- employ the plan to collect the data

Analyze Data
- select appropriate graphical or numerical methods
- use these methods to analyze the data

Interpret Results
- interpret the analysis taking into account the scope of inference based on the data collection design
- relate the interpretation to the original question

The Role of Variability in the Problem Solving Process

The *Framework* stresses the importance of understanding the role that variability plays in each component of the problem solving process.

Formulate Questions
Anticipating Variability - Making the statistics question distinction
The proper formulation of a statistics question depends on the distinction between a question which anticipates a deterministic answer and a question which anticipates an answer based on data which vary.

Collect Data
Acknowledging Variability - Designing for differences
Data collection designs must acknowledge variability in data and are designed to reduce these differences. Random selection/assignment may be imposed in order to address statistical significance.

Analyze Data
Accounting of Variability - Using Distributions
The main purpose of statistical analysis is to give an accounting of the variability in the data; the use of distributions to model variability is the key idea in the analysis of data.

Interpret Results
Allowing for Variability - Looking beyond the data
Statistical interpretations are made in the presence of deviations among the data and looking beyond the data to make generalizations must allow for this variability.

The Framework Model

The mature statistician understands the role of variability in the statistical problem solving process. The beginning student cannot be expected to make all of these linkages. They require years of experience as well as training. Statistical education should be viewed as a developmental process. The *Framework* uses three developmental levels, A, B, and C. Although these three levels may parallel grade levels, they are based on development, not age. The conceptual structure for statistics education is provided in the two-dimensional model shown in Table 1. One dimension is defined by the problem solving process components plus the nature of the variability considered and how we focus on variability. The second dimension is comprised of the three developmental levels.

Table 1: Two-dimensional model of conceptual structure

Process Component	Level A	Level B	Level C
Formulate Question	Beginning awareness of the *statistics question distinction*	Increased awareness of the *statistics question distinction*	Students can make the *statistics question distinction*
Collect Data	Do not yet *design for differences*	Awareness of *design for differences*	Students make *designs for differences*
Analyze Data	*Use* particular properties of *distributions* in context of specific example	Learn to *use* particular properties of *distributions* as tools of analysis	Understand and *use distributions* in analysis as a global concept
Interpret Results	Do not *look beyond the data*	Acknowledge that *looking beyond the data* is feasible	Able to *look beyond the data* in some contexts

Nature of Variability	Measurement variability Natural variability Induced variability	Sampling variability	Chance variability
Focus on Variability	Variability within a group	Variability within a group and variability between groups Co-variability	Variability in model fitting

SUMMARY

The complete *Framework* report is in two parts. The first part provides an overview of the goals and need for statistics education, a thorough discussion of the differences between statistics and mathematics, and a detailed presentation of the *Framework* Model. The second part is an appendix that presents descriptions of learning activities at each of three developmental levels. These provide illustrations for teachers of the types of activities that should be used in the classroom to effectively promote statistical literacy. Additionally, Franklin and Kader (2006) give a detailed example from the *Framework* that illustrates how a single problem can be used across the three developmental levels. This example illustrates how statistical thinking evolves over the three levels and includes transitional concepts between the different levels.

A good deal of progress has been made in statistics education in recent years, but there is still plenty of room for improvement. State and national standards and assessments are quite varied and the descriptions for statistics and data analysis are often poorly structured. Textbooks and other teaching materials tend to be unfocused and often include errors (unless these materials have statistics educators as part of the writing team). It is hoped that this report will provide stakeholders such as writers of national and state standards, writers of assessment items, educators at teacher preparation programs, curriculum directors, and Pre K-12 teachers with guidance in developing standards and assessments in data analysis. In the United States, the

Framework has already made a positive impact in the state of Georgia with the current revisions of the Georgia state mathematical performance standards.

REFERENCES

Cobb, G. and Moore, D. (2000). Statistics and mathematics: Tension and cooperation. *American Mathematical Monthly*, August-September, 615-630.

A Curriculum Framework for Pre K-12 Statistics Education. (2005). The American Statistical Association, http://www.amstat.org/education/gaise/GAISEPreK-12.htm.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.

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

Franklin, C. and Kader, G. (2006). A sequence of activities for developing statistical concepts. *Statistics Teachers Network*, 68, 1-12, http://www.amstat.org/education/stn/.