

## Statistics and Probability, High School

There are two unique, challenging aspects of teacher preparation in the area of statistics and probability. First, the substantially increased status of statistics in the Common Core State Standards means that most future high school mathematics teachers need to master material they never saw in their own high school education. Second, a large part of the new CCSS statistics component involves empirical, qualitative issues related to experimental design, data collection and interpretation. Realistic data analysis only makes sense in a context, e.g., one cannot talk meaningfully about stock market data without knowing something about economic events that influenced the market. This is a very different type of knowledge being taught to students than that in the standard mathematical framework of concepts, skills and problem-solving techniques in other areas of high school mathematics. Furthermore, most college statistics courses that math majors encounter devote little time to experimental design, data collection and interpretation. Thus, major changes are needed in the pre-service and in-service statistical education of high school mathematics teachers.

The Common Core State Standards for high school statistics and probability reflect the greater role for collecting and interpreting data:

### *A. Interpreting Categorical and Quantitative Data*

- Summarize, represent, and interpret data on a single count or measurement variable
- Summarize, represent, and interpret data on two categorical and quantitative variables
- Interpret linear models

### *B. Making Inferences and Justifying Conclusions*

- Understand and evaluate random processes underlying statistical experiments
- Make inferences, justify conclusions in sample surveys, experiments and observational studies

### *C. Conditional Probability and the Rules of Probability*

- Understand independence and conditional probability and use them to interpret data
- Compute probabilities of compound events in a uniform probability model

### *D. Using Probability to Make Decisions*

- Calculate expected values and use them to solve problems
- Use probability to evaluate outcomes of decisions

**Statistics.** Through pre-service and in-service education, high school mathematics teachers need to acquire the knowledge to offer good statistics and probability instruction that meets the above standards. They also should know the material in the Introduction to Statistics college course and AP Statistics course, to prepare their students for this popular college course.

Basic techniques of representing data in Standard A above, such as box plots and stem-and-leaf plots, are now frequently covered in high school and college statistics instruction, but the topics in Standard B are rarely addressed at all. For example, one important item in Standard B is, “recognize the purposes of, and differences among, sample surveys, experiments, and observational studies; explain how randomization relates to each.” The

philosophy driving this part of the statistics standards is that before students learn standard statistics inference procedures such as a t-test, they should first understand the range of issues and assumptions about where data come from and why it is variable. Variability can come just from inherent variability in the differences of random samples taken from the same population (this is the focus of traditional mathematical statistics) but it can also arise from variability induced by the experimenter, such as error-prone measurements. This in turn leads to questions about how best to collect the data and how best to analyze it once collected.

As noted earlier, this is very different instruction than is normally associated with mathematics standards. Moreover, many CCSS statistics standards have almost nothing in common with the standard upper-level probability-statistics survey course for future teachers and other mathematics majors (or even worse with the all-probability first course in a two-semester probability-statistics sequence). *The message from statistics educators is that the current probability-statistics course in the mathematics major falls well short of what future high school mathematics teachers need.*

The American Statistical Association (ASA) and NCTM have produced helpful materials to facilitate the design of a more appropriate course for future teachers: ASA's Guidance for Assessment and Instruction in Statistics Education, and NCTM's Navigation Books in Data Analysis and Probability and Essential Understandings for High School Volume in probability and statistics. A few institutions offer such courses and similar in-service courses; see the CBMS MET website for information about some of these courses.

Currently, the general-audience Introduction to Statistics course is closer to what teachers need than the upper-level math major course. However, the typical introduction to statistics course gives minimal attention to the new issues about designing studies to produce good data for answering well-formed questions. There are some recent textbooks for this intro course that do look at these issues<sup>1</sup> and would make such intro courses much more valuable for future teachers (and other students in such courses).

Given the importance of statistics in most people's daily life, given future teachers' minimal exposure in high school currently to the CCSS statistics topics, and given the holes that remain after the typical introductory college statistics course (huge holes in the case of an upper-division probability-statistics course), there is a strong argument that future teachers merit their own upper stat-prob course, using one of the introductory textbooks mentioned above. The following are the sort of questions that would be discussed.

- What exactly is wrong with saying, "I accept the null hypothesis"?
- Does a high correlation indicate that a linear model is appropriate?
- Why do different calculators and statistical software sometimes give different values for the median and quartiles?

---

<sup>1</sup> Examples are the books by Agresti and Franklin, by Peck, Olsen and Devore, by Watkins Scheaffer and Cobb, and by Starnes, Yates and Moore.

- Describe three things you could say to a high school student who asks why we use the standard deviation to measure spread.
- How would you explain why we divide by  $n - 1$  in the denominator of the standard deviation?
- What happens to the mean and standard deviation if you transform data by adding the same constant to each value? Prove your answer.
- Why is the coverage for a standard 95% confidence interval for a proportion less than 95% and what alternatives are available?
- What's the difference between the idea of independent events that we saw in the probability chapter and the idea of independence in this chi-square chapter?

In the ideal world, this would be a two-semester course. Note that this course would also be valuable to many math majors. In the absence of such a one- or -two-semester course for teachers, it would be valuable for future teachers to take, along with the intro course, a data analysis (statistical methods) course, either a general-audience sequel to the intro course or a stand-alone, upper-level data analysis course for math majors.

*Professional Development:* From the preceding discussion, it is clear that most high school mathematics teachers will need extensive professional development in statistics. In the immediate future, they will typically graduate with one upper-level probability-statistics course that covers at best a modest amount of the CCSS statistics standards. The American Statistical Association (ASA) and NCTM publication mentioned above can be used for the in-service statistical education of teachers. The footnoted textbooks for introductory statistics courses are also useful. The needed material is not technically advanced. However, because many of the statistics topics are more science than mathematics, e.g., experimentation, data exploration, hypothesis generation, it is valuable for teachers to receive some instruction from a statistician in a graduate course or some sort of mini-course. See the CBMS MET about such mini-courses and graduate courses (which could be co-listed as an undergraduate course for pre-service teachers).

**Probability.** The probability component of this area gets some coverage in current high school curricula and is reasonably covered in the upper-level probability-statistics course that future high school mathematics teachers are normally required to take in their mathematics major programs. Key topics are sample spaces, simple distributions, probability of compound events, independence, conditional probability and expectations. Note that the given greater amount of statistics that needs to be covered in a probability-statistics course, combinatorial probability problems should be minimized and left to a discrete mathematics course.

Teaching probability well is a skill that requires years of continual learning through experience. Probability is famous for its tricky logic. An example is, in an equiprobable sample space, conditioning questions can make some outcomes more likely than others—for example, if a family is known to have two children, how does knowing that one child is a girl affect the probability that the other child is a girl. Another challenge is that there are many different ways to approach some problems. Throughout their careers, teachers will

have to respond to a student's suggested solution strategy that the teacher has not seen before, possibly the strategy is correct, possibly seriously flawed, and sometimes wrong but correctible in an interesting way.

*Professional Development:* As noted above, statistics educators recommend eliminating the required upper-level probability-oriented course that future teachers take along with other mathematics majors and replacing it with a lower-level introduction to statistics or better a separate course for teachers. This will leave future teachers with a deficit in their probability knowledge. The professional development materials of the American Statistical Association and NCTM cited in the statistics section have the needed probability discussions. As with statistics, these materials can be used for: individual teachers to study on their own; to support study groups among teachers; and for use in professional development workshops. Given the subject's subtleties, these materials can be helpful for teachers who did have a probability course in college.

**AP Statistics.** The AP Statistics test was started in 1997 and now has close to 150,000 students taking it annually in over 6,000 high schools. Given the shortcomings described here in the typical pre-service statistics education of high school mathematics teachers, it is clear that extensive additional preparation in statistics is required to teach AP Statistics. Several graduate courses in statistics are desirable (chosen in individual consultation with faculty in a graduate statistics program). The minimum preparation would be a good lower-level introductory statistics course, based on the sort of textbooks mentioned above, followed by either a second undergraduate statistics course or a graduate statistics course designed for teachers (see the MET Professional Development website for details about such a course).