The Role of Dynamic Interactive Technology in Teaching and Learning Statistics

Gail Burrill, Michigan State University, USA burrill@msu.edu

Abstract
Dynamic interactive technology brings new opportunities for helping students learn central statistical concepts. Research and classroom experience can be help identify concepts with which students struggle, and an "action-consequence" pre-made technology document can engage students in exploring these concepts. With the right questions, students can begin to make connections among their background in mathematics, foundational ideas that undergrid statistics and the relationship these ideas. The ultimate goal is to have students think deeply about simple and basic statistical ideas so they can see how they lead to reasoning and sense making about data and about making decisions about characteristics of a population from a sample.Technology has a critical role in teaching and learning statistics, enabling students to use real data in investigations, to model complex situations based on data, to visualize relationships using different representations, to move beyond calculations to interpreting statistical processes such as confidence intervals and correlation, and to generate simulations to investigate a variety of problems including laying a foundation for inference. Thus, graphing calculators, spreadsheets, and interactive dynamic software can all be thought of as tools for statistical sense making in the service of developing understanding.

NEW OPPORTUNITIES

Dynamic interactive technology has the potential to extend this tool to help students understand central statistical concepts. The ability to link representations, where changes in one representation are reflected in the others, enables students to take an action, immediately see the consequences and reflect on the meaning of these consequences to make sense of the statistics- an action-consequence principle (Dick & Burrill, 2006). To maximize this potential and allow students to explore statistical concepts in deeper ways, it is possible to impose constraints on what they can do, in essence creating action-consequence "microworlds" in which students can play with a statistical concept in a variety of ways but where the opportunity to go astray, both mathematically and operationally, is limited. An action-consequence document is similar to an applet (e.g. Rice University, www.bbnschool.org/us/math/ap_stats/applets/applets.html) but can be modified or adapted by the user in ways not always possible in applets.

An action-consequence document is a technology document or file that

- requires very little knowledge on the part of the user of the device itself and how it operates;
- focuses on a fundamental statistical concept in a simple and straightforward way with both mathematical and statistical fidelity (is mathematically and statistically sound and accurate); and maintains pedagogical fidelity (does not present obstacles such as cluttered screens or too many decimal places that interfere with learning) (Dick & Burrill, 2006);
- is based on the action-consequence principle;
- have one object such as a point or graph serving as a driver for the interaction with little or no use of menus.

The design of technology-based activities for learning statistics needs careful consideration, however. There is a real danger that such materials will fall into categories absent any emphasis on what statistical learning they will enable. For example, materials such as the following are likely to appear (adapted from Belfort & Guimaraes, 2004): 1) the author's interest is on mastering the use of the technology where the statistics is secondary; 2) the activity is merely a demonstration of an idea where students are treated as spectators; 3) the activity revisits a topic to show how it can be done in a simple way with the new technology where the students' role is verification; 4) the activity replicates activities from the point of current instructional materials, underestimating the technology's potential, where the ideas are fragmented and obtaining a formula is often the objective. Materials developed for microworld environments can suffer from these same pitfalls: construction of the mathematical or statistical objects involved in the problem becomes the focus along with all of the details needed to master the technology; elaborate constructions that demonstrate a relationship but allow no interaction on the part of students, use of the technology to perform relatively meaningless operations (i.e., entering 100 numbers to find the summary statistics); replicating what can be done on a graphing calculator with no attention to the opportunities afforded by the new capabilities.

To exploit a dynamic environment, just as with handhelds and dynamic statistical software (Doerr & Zangor, 2000; Schwarz & Hershkowitz, 1999) students need to have adequate opportunities to conjecture, reflect, explain, and justify. Thus, technology's influence on students' statistical learning is either amplified or limited through the kinds of statistical tasks teachers provide and the questions they ask. Thinking and conjecturing on the part of the students can be enhanced or inhibited depending on the kind of answers the questions elicit. Questions such as "What did you get?" or "What's the next step?" will not prompt student thinking as much as "What does this mean?" and "How is this different?" - questions that push students past passive recording to active sense making.

The discussion below examines how these dynamic interactive worlds together with questions and tasks focused on reasoning and making sense of the statistics might address issues related to understanding central statistical concepts using several examples from analyzing single variable data and building understanding of inference.

VARIATION

In general, the research makes clear that students struggle with the concept of distributions, failing to abstract from individual data points to view the collection as a whole with its own characteristics. Students should be able to integrate both centers and spreads when they deal with data, yet Cooper and Sloan suggest that student understanding of variability is "tenuous" (2008). Students often describe distributions using language such as "the dots are close together here and spread out there", which Bakker (2004) calls "local views" on spread. In his study, none of the students viewed spread as a dispersion from a mean or median value. Many texts introduce the concept of mean and median as algorithmic processes, without first exploring the notion of center, and very few connect either of these to the concept of spread.

Standard Deviation.

Using a dynamic interactive plot, students can explore distributions and, in particular, the concept of standard deviation as a measure of variation. Allowing students to experiment with a distribution, noting how the data cluster around the mean can precede the formal definition of standard deviation and can help them reason about when such a measure would be sense making and useful. When the user grabs and drags a point or set of points (figure 1), the distribution changes, as do the corresponding values of the mean, and the plus/minus interval (one standard deviation from the mean).

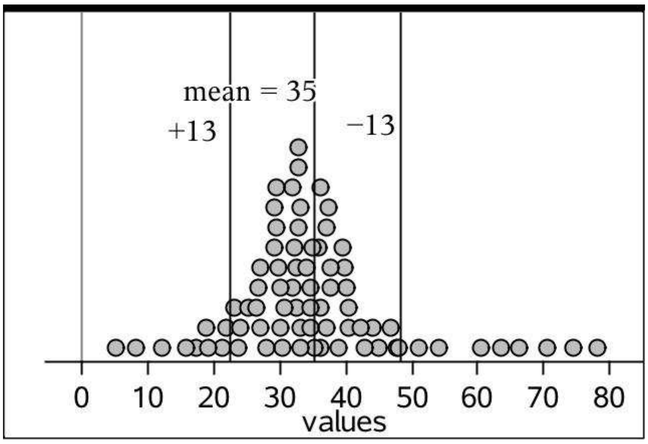


Figure 1: Mean and standard deviation

Research suggests interventions that support the development of aggregate reasoning in younger students are more successful if they are rich in student discussion of qualitative aspects of distribution such as shape, including identifications of the location of hills or deviation bumps, gaps, and spread-out-ness (Konold & Kazak, 2008). To promote such discussions, students can be asked to move points to find a distribution with certain characteristics, such as mean of 35 and a larger interval than 19 or find a distribution where the plus/minus interval would not convey much meaningful information about how the points were distributed. Conversations that encourage students to describe their distributions can help them develop a sense of what constitutes a distribution. Students can be given a mean and an interval and asked to create distributions with those characteristics, comparing their work and observing what the distributions have in common. Questions such as "What if…" or "Is it possible to…. Why or why not?" can help students focus on the relationship between the mean and the interval as a characteristic of the distribution. Asking students to describe possible contexts for distributions they create and to interpret the mean and interval in terms of those contexts can help them make the connections that lead to understanding. The goal of the activities should be to help students develop a sense of an expected variability that would seem to be reasonable around the expected value, almost like an intuitive confidence interval (Shaughnessy, 2006).

Box Plots Box plots are a powerful tool for comparing distributions, but research has shown that some of their features make them particularly difficult for students to interpret. These difficulties include: individual cases are usually hidden in box plots; box plots operate differently than other displays (the area of a histogram represents the frequency, for example); the median is not intuitive to students who often think of it as a cut point not as a measure of center and thus as a characteristic of the data; quartiles divide the data into groups in ways that are difficult to understand (Bakker, Biehler, Konold, 2004). Action-consequence documents are promising resources that might be used to address some of these difficulties and to better develop students' ability to interpret box plots.

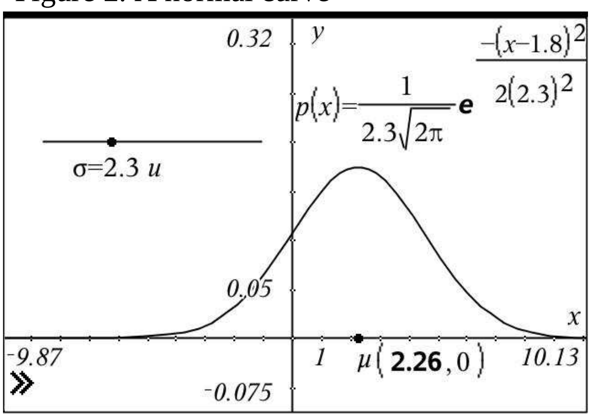
Consider a set of possible times it might to take for students to get to school displayed in a document where details such as the same scales for the dot plot and the box plot and the presence of the mean in the dot plot are not left to chance, and students can focus on the statistical concept rather than on creating the document. Selecting a data point (or points) highlights the position of that point in the box plot. Grabbing and moving a point(s) changes the value of the point(s) and thus, the distribution.

Students can highlight points inside the box plot and relate them to the corresponding points on the dot plot; teachers can ask students to describe the relationship between the group in that region of the box plot and the distribution. Students can change the distribution by moving the points, exploring questions such as, "What would the distribution look like if the median is at the lower quartile?" or "Can a distribution have a box plot with no whiskers? What would this suggest about the data?" If you drag one of the segments of the box plot, the distribution will also change, providing the opportunity for teachers to ask students about connections among the distribution of the data and the box plot, the role of the median and its relation to the data.

Inference Students enter and often leave statistics education courses with misconceptions. For example, Lunsford and colleagues (2006) found that students at the end of a course still confused variability with frequency, had problems with the “averaging reduces variation” concept and did not fully understand that for a fixed sample size, the sample mean was a random variable and thus had a distribution with a shape, center, and spread. Technology can address some of these issues, and, in addition, some studies suggest that technology greatly facilitates a “predict-and-test” strategy that can establish the cognitive dissonance necessary for students to change their ways of thinking about a concept (e.g., Posner et al, 1982).

Normal Curve How many normal curves are there? This question can produce surprising answers. Students often think there is only one normal curve, with mean 0 and standard deviation1. They have trouble understanding that just as other functions have a basic structure with the characteristics determined by varying parameters, a family of normal curves is determined by the mean and standard deviation. Investigating the relationship between the graph, the mean and the standard deviation can help students anchor and generalize their concept of a normal curve (figure 2).

Figure 2. A normal curve



To understand how a normal curve behaves, students can respond to questions such as "How is p(x) the same or different from $f(x) = ax^2+bx+c$?"; "What will happen to the distribution when the mean is changed?"; "the standard deviation?"; or "Compare changing the mean to changing the standard deviation." Answering these questions will force students to think about how the distribution connects to other mathematics they know and, if the curve is plotted on a grid, can set the stage for understanding that the area under the curve is the same for any combination of standard deviations and means.

Simulations

Research suggests that simulations can develop students' understanding by having them carry out many repetitions (i.e., comparing different samples of the same size from a population), controlling parameters (such as sample size), describe their observations and reflect on what these mean statistically rather than on concentrating on theoretical probability discussions, which can often be counterintuitive (delMas, Garfield, & Chance 1999). Interactive dynamic technology enables students to generate a sample from a given population, calculate a statistic for that sample, and look at the distribution of the statistic, all on a screen simultaneously. The repetition of many simulations for distributions using the same parameters can help address the concern students' understanding of statistical inference seems to be hindered by a limited understanding of related concepts such as distribution and variability (Chance et al, 2004). By pressing a control key, students can be asked to compare *t* distributions to *z* distributions from the same population, noting the consistent difference in shape for a given sample size and what changing the sample size does to this difference. They can investigate the central limit theorem for distributions that are normal and those that are not and how changing the sample size affects the distribution of the sample means (figures 3 and 4).

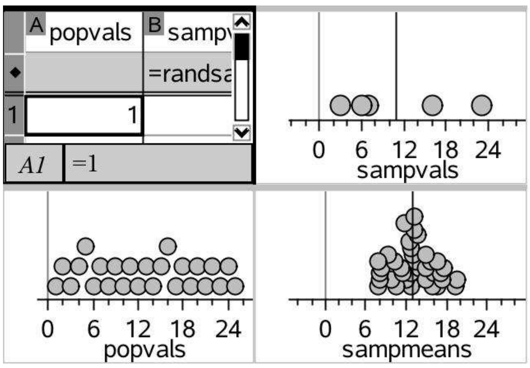


Figure 3: Sample means sample size 5

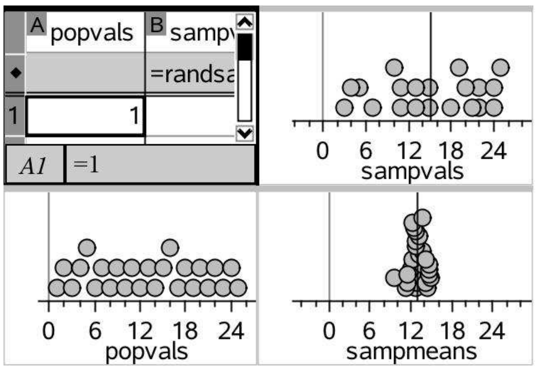


Figure 4: Sample means sample size 20

These "action consequence" documents as well as others that allow students to investigate confidence intervals, probability distributions such as the chi square or geometric, or what it means to have a significant observation can engage students in reasoning about the conditions, assumptions and theoretical tools involved in using statistics and random sampling to make inferences about some unknown aspect of a population. The questions posed can help students make connections to mathematics such as interpreting graphs they have encountered in other mathematical settings. A caution here, however, is in order to maximize the benefit of these interactive environments, the use of computer simulation for demonstration purposes only is not sufficient for developing real understanding of the concepts, in particular sampling distributions and the central limit theorem (Lunsford, Rowell □& Goodson-Espy, 2006).

CONCLUSION

Technology has clearly been instrumental in the practice of statistics where statistical tables are no long needed as calculators easily provide accurate values, "resampling statistics" (Good 2006) offers an intuitive alternative to model-based inferential models, and new methods for visualizing and exploring data are emerging as powerful tools for analysis (Chance et al, 2007)

As technology advances and more students have access to dynamic interactive technologies, more opportunities become available for helping students learn. The discussion above has been on the use of this technology to support student learning in a targeted way, where the focus begins with core concepts with which students struggle. Dynamic interactive technology allows the creation of action consequence documents that make possible not only a re-examination of "*what* [statistics] students should learn as well as *how* they can best learn it" (NCTM, 2000).

References:

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, *3*(2), 64-83. www.stat.auckland.ac.nz/~iase/serj/SERJ3(2)_Bakker.pdf

Bakker, A., Biehler, R., Konold, C. (2005). Should young students learn about box plots. In G. Burrill & M. Camden (Eds.). *Curricular Development in Statistics Education*. International Association for Statistical Education 2004 Roundtable. Voorburg, the Netherlands: International Statistical Institute.

Belfort, E. & Guimaraes, C. (2004). Teachers' practices and dynamic geometry. In Hoines, M. & Fuglestad, A. (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education.* 2, 503-510).

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 277–294). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Chance, B., Ben Zvi, D., Garfield, J., & Median, B. (2007). The role of technology in improving student learning in statistics. *Technology Innovations in Statistics*, 1(1).

Cooper, L. & Shore, F. (2008). Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. *Journal of Statistics Education* 16(2), www.amstat.org/publications/jse/v16n2/cooper.html

delMas, R. C., Garfield, J., & Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education* [Online], 7(3). http://www.amstat.org/publications/jse/secure/v7n3/delMas.cfm

Dick, T. & Burrill, G. (2006). Presentation at National Council of Supervisors of Mathematics Annual Meeting, Atlanta GA.

Doerr & Zangor (2000). Creating meaning for and with the graphing calculator. *Educational Studies in Mathematic, 41, 141-163.*

Good, P. (2006). *Resampling methods* (3rd Ed). Birkhauser.

Konold, C. & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education*, 2(1), Article 1. http://repositories.cdlib.org/uclastat/cts/tise/vol2/iss1/art1

Lunsford, M.L., Rowell, G., & Goodson-Espy, T. (2006). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. amstat.org/publications/jse/jse_archive.html

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

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211-227.

Schwarz, B. B., & Hershkowitz, R. (1999). Prototypes: Brakes or levers in learning the function concept? The role of computer tools. *Journal for Research in Mathematics Education, 30*(4), 362-389.

Shaughnessy, M. (2006). Research on students' understanding of some big concepts in statistics. In G. Burrill (Ed.) *Thinking and Reasoning with Data and Chance*: 2006 Yearbook. Reston VA: National Council of Teachers of Mathematics.