USING SIMULATION TO TEACH AND LEARN STATISTICS

Beth Chance and Allan Rossman
California Polytechnic State University, United States
arossman@calpoly.edu

Technology, and simulation in particular, can be a very powerful tool in helping students learn statistics, particularly the ideas of long-run patterns and randomness, in a concrete, interactive environment. This talk will provide examples of the integration of simulation to enhance topics throughout an introductory statistics course through a combination of Minitab *macros and specifically designed applets. Topics will include randomization tests for comparing groups, and sampling distributions of proportions, odds ratios, and regression coefficients. We will also highlight how simulation can motivate students to learn the more mathematical derivations. Feedback and sample work from students will be presented, as well as issues in designing effective simulation investigations.*

INTRODUCTION

Advances in technology have enabled heavier use of such tools by both instructors and practitioners of statistics. These developments have enhanced the ability of the practitioner to solve many problems that do not have straight forward analytic solutions, enlarging the toolbox of statisticians. Using software to directly involve students in the analysis of data is clearly a beneficial exercise, but we also argue that technology, and simulations in particular, can be a very effective instructional tool. In teaching statistics, it is important to both model the use of technology as a problem solving tool and to take advantage of the ability to automate calculations and graphics, facilitate explorations of statistical concepts through dynamic, interactive, visual environments, and experience first hand the stochastic nature of statistical processes. The impact of technology on the teaching of statistics has reached all grade levels.

One of the most challenging aspects to teaching and learning statistics is that so many statistical concepts and methods are based on the issue of what would happen if a random process (such as random sampling from a population or random assignment of subjects to treatment groups) were repeated indefinitely. This abstract notion is very difficult for most people to grasp. Technology provides the opportunity to make this abstract idea more concrete by enabling students to repeat such random processes a very large number of times and describe their observations first hand. Mills (2002) provides an overview of using computer simulation methods to teach statistics, though empirical research in statistics classrooms has been limited. Research in other fields, particularly in educational technology, has suggested several guidelines for using simulations in instruction. For example, Nickerson (1995) offers the following maxims:

- View learning as a constructive process where the task is to provide guidance that facilitates exploration and discovery.
- Use simulations to draw students' attention to aspects of a situation or problem that can easily be dismissed or not observed under normal conditions.
- Provide a supportive environment that is rich in resources, aids exploration, creates an atmosphere in which ideas can be expressed freely, and provides encouragement when students make an effort to understand.

Others (e.g., Snir, Smith, and Grosslight, 1995; Glencross, 1998) have provided additional recommendations including using simulation to make abstract concepts more concrete and correlated different representations of the same concept, while providing immediate feedback to the students in a more appealing manner. For example, delMas, Garfield, and Chance (1999) also demonstrated a powerful effect of using computer simulations in a predict-and-test environment (to create cognitive dissonance as discussed in science education literature by Posner *et al.* (1982) and others) in aiding student understanding of sampling distributions. With these recommendations in mind, we have developed several curriculum enhancements that we believe utilize simulations to effectively help students learn and do statistics. We describe five sample activities in the following and then conclude with our recommendations for using simulations in an introductory, tertiary level statistics course for non-majors. Many of these activities describe java applets that can be accessed at http://www.rossmanchance.com/applets/, although the choice of technology is less important than the design of the accompanying classroom activity.

EXAMPLE 1: PROBABILITY (Random Babies)

In this activity, students are apologetically introduced to a scenario where four mothers give birth to four baby boys at a hospital but then they are asked to suppose that the infants are returned to their mothers at random (see Rossman and Chance, 2001). The goal is for students to investigate the probability that all four mothers receive the correct babies in a way that allows them to develop an intuitive understanding of what we will mean by "probability." Students are first given index cards to simulate the distribution of the babies to the mothers (each index card is given a different baby name with another piece of paper used to represent the four mothers). This activity is then extended to a computer simulation utilizing the same context (Figure 1).

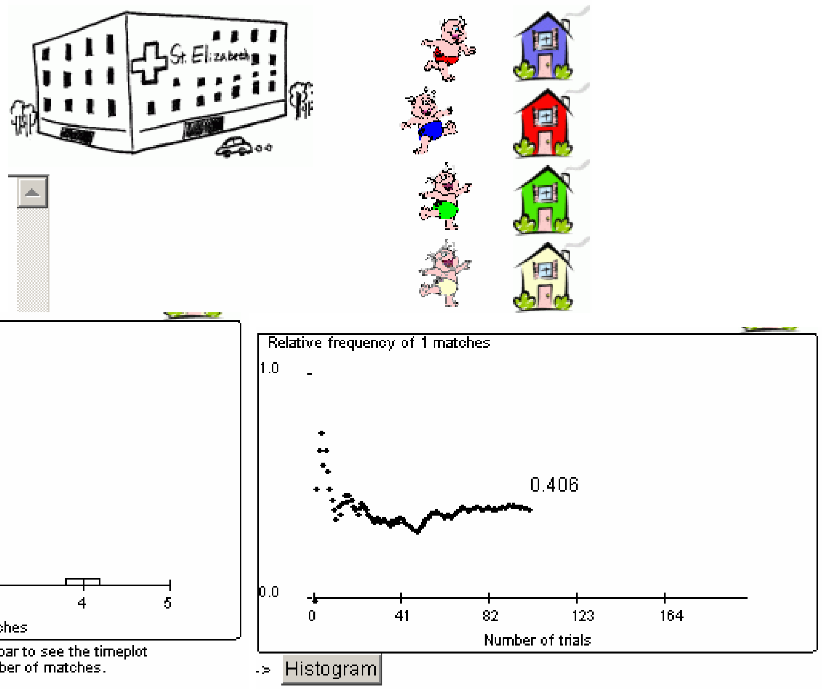

Figure 1: Babies are assigned at random. The applet cumulative counts the number of matches and a timeplot over the number of trials can be displayed.

Through this applet, students can directly mimic their hands-on simulation (reducing cognitive load). We ask them to first focus on individual random assignments and then ask them to explore the long-term properties of the number of matches as the number of trials increases. Students directly visualize the empirical probability converging to its limiting value (which can later be superimposed on the graphs). This approach is easily extended to the concept of the expected value and to larger cases (e.g., n = 8 babies) that are not as tractable analytically.

EXAMPLE 2: RANDOM SAMPLING (Gettysburg Address)

In this activity, students are asked to select a convenience sample of ten words from the text of Lincoln's famous Gettysburg Address (see Chance and Rossman, 2006). When then asked to explore the average length of words in their samples, students typically discover that they consistently overestimate the population mean. Students intuitively suggest selecting "random samples" instead and using the computer, they select random samples of five words and see that their sample means now provide unbiased estimates of the population mean. We again extend this more tactile simulation to a java applet that allows them to explore factors affecting this sampling distribution such as sample size and population size (see Figure 2). In conducting these simulations, students predict, discover, and then attempt to explain the behavior they observe in the shape, center, and spread of the empirical sampling distribution.

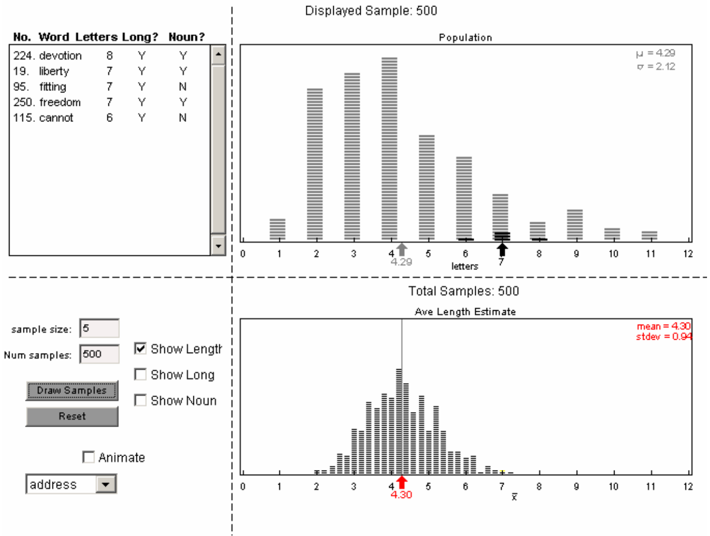

Figure 2: Population, (last) sample, and empirical sampling distribution of 500 samples of size *n* =5 of the length of words in the Gettysburg Address

This activity can also be carried out using a *Minitab*® macro that allows students to write a self-contained program to repeatedly invoke the same set of commands. In the same manner, students can easily explore the effects of different factors such as the sample size, population shape, and population size (while also paying close attention to the distinction between the sample size and the number of samples). An advantage to this approach is that it is easily extended to explore the distributions of other statistics and the idea of bootstrapping in general. A disadvantage is that students lose the more direct link to the contextual basis for the exploration and the learning curve of the technology may be longer. However, by utilizing the simulation approach as a theme throughout the course, students begin to focus on the larger, constant question of "What would happen if we took many random samples?"

EXAMPLE 3: SAMPLING DISTRIBUTION OF REGRESSION LINES (Studying and Grades)

In this activity, students are given data from a previous student project that examined self-reported grade point averages (GPA) and weekly study time for a sample of 80 college students (see Rossman and Chance, 2001). They examine the scatterplot, correlation coefficient and least-squares line (Figure 3).

Students are then asked informally if they think the observed relationship between these two variables could have happened "by chance." By this point in the course, they are ready to consider how often a sample slope at least this extreme would occur from random sampling if the population slope was zero, and they can use technology to explore the sampling distribution of the sample slopes.

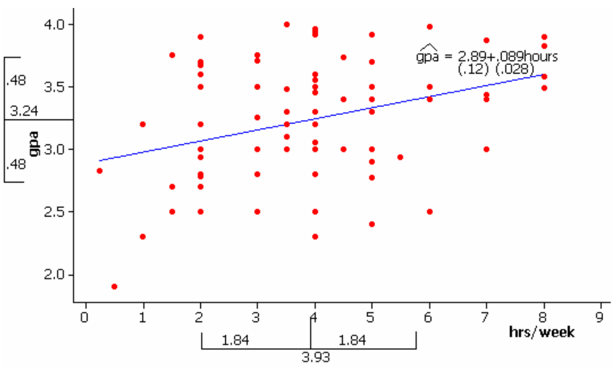

Figure 3: Sample distribution of student self-reported GPA and study time per week

Students are presented with a large population with zero correlation between GPA and study hours but otherwise matching the descriptive statistics of the original project data (Figure 4). Students can then select a random sample of 80 individuals from this population and view the resulting sample regression line.

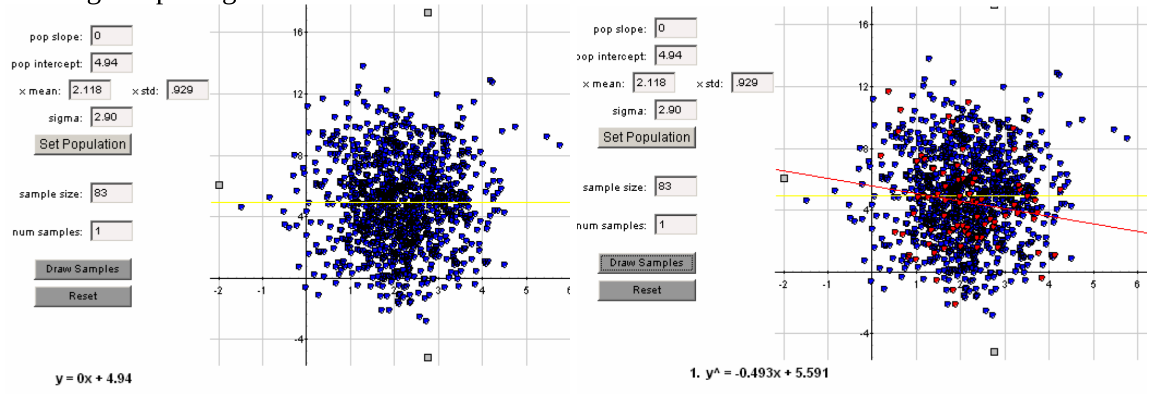

Figure 4: Bivariate population (left) with one sample selected in red (right)

After understanding this process, students are asked to draw a large number of samples and view the empirical sampling distribution of the sample slopes (Figure 5). Students can also view a display of the generated sample regression lines.

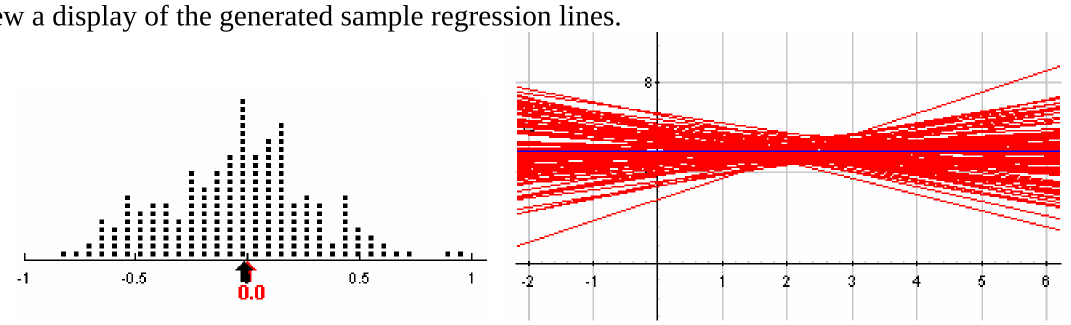

Figure 5: Empirical sampling distribution of sample slopes and fifty sample regression lines

At this point, students should not be surprised by the symmetric distribution centered at the value of the population parameter that also decreases in variability if the sample size is increased. In fact, students can be led to explore the components of the formula for the standard error of the sample slope (sample size, standard deviation of explanatory variable, and variability about the regression line) as well as the "bow-tie" shape of prediction and confidence intervals.

EXAMPLE 4: SAMPLING DISTRIBUTION OF ODDS RATIOS (Sleep deprivation)

In this activity, we present students with data from a case-control study (Conner *et al.*, 2002) in which researchers compared sleeping habits of a case group that had been drivers in a car crash and a control group that had not been in a crash (Chance and Rossman, 2006). Preliminary discussion justifies using the odds ratio as the summary statistic because of the case-control nature of the data collection. This motivates students to need to know the sampling distribution behavior of this statistic. Building on what they have done throughout the course, students turn to simulation to investigate the behavior of sample odds ratios under the null hypotheses of no difference (in proportions who had been sleep deprived) between the two binomial populations.

Students immediately encounter a surprise with this simulation: rather than the typically symmetric behavior they have been seeing, the empirical sampling distribution of the sample odds ratio turns out to be positively skewed (Figure 6). This motivates students to try a transformation. The empirical sampling distribution of the sample log odds ratio looks quite symmetric and normal (Figure 6) and students justify applying normal-based tests and confidence intervals with the sample log odds ratio.

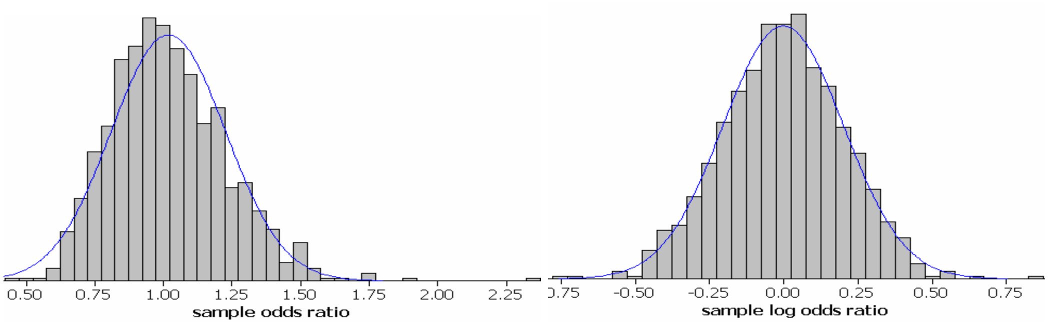

Figure 6: Empirical sampling distributions for sample odds ratio and log(odds ratio)

EXAMPLE 5: SIMULATING RANDOMIZATION TESTS (Sleep deprivation)

Through simulations, students can not only explore fundamental ideas such as statistical significance but also employ more modern statistical techniques such as empirical randomization tests. In this activity, students are presented with results from a randomized experiment of the effects of sleep deprivation on the visual discrimination performance of 23 subjects three days later (Stickgold *et al*., 2000). Students produce a *Minitab*® macro to simulate repeated random assignments of the 23 subjects' improvement scores to the two treatment groups, and then explore the empirical distribution of the difference in group means (Chance and Rossman, 2006; Figure 6). Students calculate an empirical p-value by counting how many of the simulated randomizations produce a difference in group means as extreme as in the actual study.

```
random 23 c2 c3
unstack c1 c4 c5;
subs c3.
let c6(k1)=mean(c4)-mean(c5)
let k1=k1+1
```

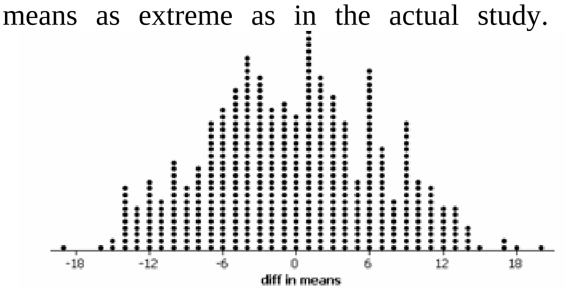

Figure 7: Macro and example results of empirical randomization test

We believe that this simulation of a randomization test can lead students to a deeper understanding of the concepts of statistical significance and p-value than, say, conducting *t*-tests. Moreover, students are learning a modern, flexible, computer-intensive analysis technique in the process. Cobb (2005) goes so far as to suggest that the entire introductory statistics curriculum could be refocused on randomization tests and bootstrapping, with simulation as the analysis tool.

IMPLEMENTATION SUGGESTIONS

In using simulations to help introductory statistics students explore such statistical concepts, we present the following recommendations:

- Start with a hands-on simulation and directly involve students in the simulation process (this provides them with ownership as well as a better understanding of the process so that transition to a computer simulation does not become a black-box demonstration).
- Create a link between the tactile and computer simulations. Such a link can be established by simply embedded the technology simulation in the same context. Distractions should be minimized and the technology should never appear as a "black box" to the students. Design the simulations to match the randomness of the study design (e.g., random sampling vs. random assignment).
- Choose technology to facilitate student interaction and accessibility~~,~~ maintaining the focus on the statistical concept rather than on the technology. This choice can depend on learning curve, portability, and whether the technology tool can be utilized in other places in the course (e.g., using the same software to carry out traditional data analysis tasks). It is important to consider the background of the students and the goals of the course as well as instructor comfort level and knowledge.

- Use tools that allow quick, immediate, and visual feedback. Java applets and other "click and drag" software (e.g., *Fathom*®) appear especially useful for these features. In particular, a predict-and-test model can be very effective in establishing cognitive dissonance so that students will pay more attention to their own understanding, and correction of misunderstandings, of the concept.
- Carefully design the learning activity to guide student interaction. Set-up time should be minimal and students will, at least initially, need to be carefully stepped through the activity, which the steps building logically but also always encouraging them to focus on the overall larger statistical concept. Follow-up instruction (e.g., debriefing, student assessment) should always be provided *after* the students have initially explored the concept.
- Build on the simulations throughout the course enabling students to see simulation as an analysis tool in its own right, while also reinforcing key ideas (and reducing cognitive load). Students should consider using simulations to compare different analysis techniques (e.g., evaluate the performance of different confidence interval procedures) and become comfortable with the idea of an empirical p-value.

CONCLUSION

We contend that simulations can be extremely effective learning tools for helping students to understand abstract concepts associated with repeated random processes. Standard software packages can be useful as can specially designed software such as applets. The key is to start with the concept that needs to be addressed and to carefully design the learning activity to guide student interaction and help them focus on the key ideas. More empirical studies need to be conducted to help pinpoint the most effective components of using simulations as a pedagogical tool.

REFERENCES

Chance, B. and Rossman, A. (2006). *Investigating Statistical Concepts, Applications, and Methods,* Belmont, CA: Duxbury.

Cobb, G. (2005). The introductory statistics course: A saber tooth curriculum? Talk presented at United States Conference on Teaching of Statistics, Columbus, OH.

Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., Bailey, J. and Jackson, R. (2002). Driver sleepiness and risk of serious injury to car occupants: Population based case control study. *British Medical Journal*, 324, 1125.

delMas, R., Garfield, J. and Chance, B. (1999). Assessing the effects of a computer microworld on statistical reasoning. *Journal of Statistics Education*, 7(3), http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm.

Glencross, M. J. (1988). A practical approach to the Central Limit Theorem. In R. Davidson and J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*, Victoria, B.C.: University of Victoria, (pp. 91-95). Voorburg, The Netherlands: International Statistical Institute.

Mills, J. D. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education*, 10(1), http://www.amstat.org/publications/jse/v10n1/mills.html

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

Rossman, A. and Chance, B. (2001). *Workshop Statistics* (2nd edition). Emeryville, CA: Key College Publishing.

Snir, J., Smith, C., and Grosslight, L. (1995). Conceptually enhanced simulations: A computer tool for science teaching. In D. N. Perkins, J. L. Schwartz, M. M. West, and M. S. Wiske (Eds.), *Software Goes to School: Teaching for Understanding with New Technologies*, (pp. 106–129). New York: Oxford University Press.

Stickgold, R., James, L., and Hobson, J.A. (2000). Visual discrimination learning requires post-training sleep. *Nature Neuroscience*, 2, 1237-1238.