ASSESSMENT OF THE PEDAGOGICAL UTILIZATION OF THE STATISTICS ONLINE COMPUTATIONAL RESOURCE IN INTRODUCTORY PROBABILITY COURSES: A QUASI-EXPERIMENT

Ivo Dinov and Juana Sanchez
University of California - Los Angeles, United States
jsanchez@stat.ucla.edu

Java-applets have proved to be effective in different settings to demonstrate statistical concepts. The NSF-funded Statistics Online Computational Resource (SOCR) provides a number of interactive tools for enhancing instruction in various undergraduate and graduate courses in probability and statistics. In this paper, we present the results of a quasi-experiment aimed at assessing the effect of moderate use of SOCR on three outcome measures: course scores, student satisfaction and choice of technology to complete the final examination. One section of an upper division "Introduction to probability" class, the treatment group, was conducted with the instructor using the probability applets and materials as assessment tools in homework. Another section, the control group, of the same class was conducted exactly the same way, but not using SOCR. We compared the three outcome measures in the two groups and discuss ways to improve the quasi-experiment. Ideas are given on how to incorporate the applets in various probability courses.

STATISTICS ONLINE COMPUTATIONAL RESOURCE (SOCR)

The Statistics Online Computational Resource (SOCR) provides a number of interactive tools for enhancing instruction in various undergraduate and graduate courses in probability and statistics. These resources include online applets, course materials and activities. The suite of web applets includes tools for concept demonstration, intuition building, statistics and probability computing, data analysis, model fitting and virtual experimentation. The course materials include tutorials, lecture notes, instructor's aids and problems. SOCR provides an easy and extensible framework for contributing new plug-in tools within this dynamic interactive environment.

We describe in this paper a quasi-experiment where we use the following modules of SOCR in an *Introduction to Probability* class: *Interactive Distributions* (http://socr.stat.ucla.edu/htmls/SOCR_Distributions.html) and *Experiments and Demos* (http://socr.stat.ucla.edu/htmls/SOCR_Experiments.html). The first module allows interactive computation of probabilities for discrete and continuous random variables, using a graphical interface that allows students to see pictorially what they are calculating. The second module contains a number of experiments, such as the birthday experiment, the matching experiment, central limit theorem experiments, bivariate normal experiments, etc. To accompany these tools the instructor of the courses (JS) prepared handouts with instructions to teach students how to use the applets in a basic way and other handouts with activities that required the use of the applets to discover the answers to more complex applied probability problems. The use of the applets to complete these activities was required and had to be turned in as homework for a grade. Homework was worth 20% of the course grade. After this training, students were allowed to use the applets for the final exam, but this was not required. At UCLA, computing support is available to all students in the form of campus labs and laptop checkout which can last from 3 hours to the whole day depending on demand. The SOCR activities can be found at http://www.socr.ucla.edu.

WHAT DO WE KNOW ABOUT APPLETS' EFFECTIVENESS IN TEACHING PROBABILITY?

Studies have found that students in introductory statistics courses react very positively to applets and interactive aids, both in term of enjoying playing with them as well as better understanding the concepts (Anderson-Cook *et al*., 2003; Aberson *et al*., 2002; West and Ogden, 1998). Because of lack of access to technology for all students, or for other reasons, no testing of the effectiveness of using the applets in student performance or other outcome measures exists. Even less evidence exists on the effectiveness of applets in teaching probability courses.

Experts agree that any teaching of probability should address several dimensions (Metz, 1997): (1) Understanding what random and chance variation means. (2) Being able to interpret a probability distribution and make well reasoned claims about a variable by studying its probability distribution. (3) Have students be able to interpret a display of a probability distribution and understand how it conveys probability. (4) Know how to work with a probability distribution and see in it the difficult concept of probability. In a course dedicated exclusively to the teaching of Probability, there are other concepts such as different roles of marginal and conditional probabilities, probabilities of functions of random variables, moment generating functions, approximation theorems and central limit theorems that are very important in the curriculum. However, when students have never taken a Statistics course prior to taking the mentioned probability course, the dimensions mentioned above are as relevant as in an Introduction to Statistics course. Probability distributions are crucial for learning almost any part of probability and statistics (Cohen and Chechile, 1997). SOCR experiments are particularly suited to this task, since students generate data and they can compare their data distribution with the theoretical distribution at a very early stage. The applets from SOCR that we use in the quasi-experiment described in this paper allow developing all those dimensions mentioned above. The assessment instruments used (homework, midterms, final exam) all tested student understanding of the role of probability distributions.

QUASI-EXPERIMENT: EFFECTS OF SOCR ON STUDENT LEARNING, SATISFACTION AND USE OF TECHNOLOGY

To determine the effectiveness of the use of SOCR tools in student learning, student satisfaction and use of technology, we conducted a quasi-experiment with two Introduction to Probability classes that were as similar as possible in all characteristics relevant to the outcome measures of interest. The only difference was that in one group we used SOCR and in the other group we did not. This Introduction to Probability class is, for many students, the first exposure to Statistics that they have. The requirement is univariate and multivariate calculus only.

The two classes participating in the experiment were taught during the Fall quarter 2005. The treatment (SOCR) group, Section 4, received instruction at 9:00 am Monday, Wednesday and Friday with a discussion session conducted by a teaching assistant on Thursday. The control group, Section 3, received instruction at 11:00 am the same days of the week and a discussion session with the teaching assistant on Tuesday. Both classes were taught in a room with data projector and using the same power point lectures, which were also available in the course web site; both classes had the same core homework and similar discussion sessions. The treatment group had additional activities in the homework that required SOCR and some of the common core problems were required to be done in SOCR. Figure 1 contains the academic composition of the two groups with all students included.

Group	*Major (%)*						*Class (%)*			
Treatment (*n*=20) 9-9:50am	Math	0.15	Ap Math	0.05	Math/Ec	0.35	Junior	0.65	Senior	0.151
	Math-app	0.25	Anth(G)	0.05	Gen(G)	0.05	GDI	0.15	Other	0.05
	AtSci(G)	0.05	Unex	0.05						
Control (*n*=39) 11-11:50am	Math	0.051	ApMath	0.025	Math/Ec	0.205	Junior	0.28	Senior	0.28
	Math-app	0.051	PrBEco	0.025	Biost(G)	0.33	GDI	0.025	GD2	0.051
	Mph(G)	0.051	Bioch	0.025	Unex	0.025	GMT	0.33	Other	0.025

Figure 1: Composition of the two groups with all students included

Treatment group students were also taught how to use *R*, but *R* was never required. The control group did not have those activities, but in some homework they were required to use the software *R*. The teaching assistants conducted the activities that I would assign to them, including training students to use *R* and the applets, and would leave some time to answer homework questions. Based on the summaries that the Teaching Assistants wrote for us each week, the two sessions were conducted similarly. Both classes had identical information in the syllabus, identical practice midterms and final, and identical information about labs and accessories, but only the treatment group, session 4, had a link to SOCR in the web page. Because of password

requirements and different teaching assistants and due date for different homework, the students in one section were not interacting with the students in the other. The textbook use in both classes was Ross (2005). All exams were graded simultaneously, with each of us grading a few questions for all students. The reader grading the homework was different, but both readers were given the same answer key and point allocation instructions. After we started the study, we realized that the enrollment in the control group had a large number of graduate master students from the Biostatistics Department and a few others from other departments. The treatment group only had three graduate students from different departments. This created a confounding factor in the comparison between the two groups. For this reason, in Figure 2, you will see the outcome measures for the treatment and control group without graduate students in any of the groups. It is the table in Figure 2 that we discuss here. Figure 3 compares the total score.

Student Learning Outcomes

Compared with the undergraduate students in the control group, the undergraduate students in the treatment group did slightly better than the control group in all outcome measures relevant to student learning (homework, midterms, final exam and total score), as we can see in Figure 2. These results are reversed when we include the graduate students in the comparison, showing the impact of having the biostatistics students. As we can see in Figure 2, and the box plots of Figure 3, we do not have outliers among the undergraduates, and the variability in total score is much smaller in the SOCR group than in the control group.

Group	Midterm (35)	Homework (20)	Final (45)	Total (100)
Treatment *N*=17	mean=27.08 median=28.5 min=17.5 max=33 sd=4.25	mean=16.41 median=16.41 min=13.39 max=19.07 sd=1.73	mean=29.23 median=29.40 min=22.5 max=36.9 sd=4.30	mean=72.73 median=72.44 min=58.42 max=84.89 sd=8.08
Control *N*=23	mean=26.19 median=26.5 min=17.5 max=34 sd=4.65	Mean=15.76 Median=16.92 Min=4.69 Max=19.74 sd=4.44	Mean=28.48 Median=29.4 Min=18.4 Max=37.2 sd=5.31	Mean=70.45 Median=71.22 Min=43.19 Max=90.55 sd=12.24

Figure 2: Student Learning Outcomes

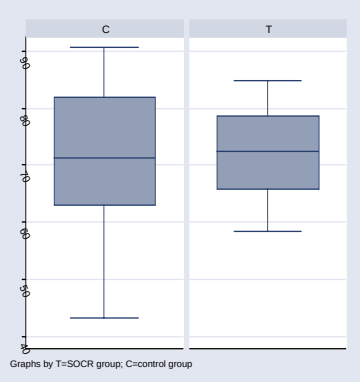

Figure 3: Total scores

Use of Technology Outcomes

The final exam for both groups was conducted in a computer lab where all the computers were centrally monitored by the Teaching Assistants while students worked. The students could have only *R* and SOCR opened (treatment group) or only *R* (control group). The Teaching Assistant would check their monitors every 15 minutes to record the activity and kept an eye on all monitors constantly to guarantee that there was no other software opened. In the treatment group, by the end of the final exam, 6 students had at some point used R, and 13 had used SOCR. Thus in the treatment group, 95% of the students used one or another form of technology, with 65% using SOCR. In the control group, by the end of the exam, about 26 (68%) of the students had used the technology.

It must be pointed out that all students were allowed to have a scientific calculator during the exam. Moreover, all questions could have been answered using only the scientific calculator and remembering some results that had been obtained with software earlier.

Student Satisfaction Outcomes

At the end of the quarter students filled out a questionnaire where they indicated the following: (a) Whether the use of technology in the class had made their learning more effective or not in some dimension, as compared with other classes not using technology (79% in the treatment group said yes to some dimension, 67% in the control group); (b) Whether they would take a second probability course with more applications (73% in the treatment group said yes,

81% in the control group); (c) Whether the class taught them things more relevant to their lives than other classes (84% in the treatment group, 62% in the control group). Most students in both groups reported that this class was more interesting and easier than other applied mathematics classes they had taken.

DISCUSSION

The study described in this paper is only a small scale experiment with moderate use of the SOCR applets. The conclusions, although encouraging, should only be taken as preliminary testing. It is not as easy as one would expect to embark the design of a probability course with SOCR. The experiments and activities require giving students specific instructions on how to use the applets, and then also giving students some more advanced activities that would require a non-trivial use of the applet. The quasi-experiment described in this paper was an initial attempt at doing so, that is why we call the treatment a "moderate" use of the applets. In order to make the instruction as similar as possible in the two groups, the applets were not extensively used in the lecture sessions but the applets used in the activities that students had to do were demonstrated in class, or discussed in the Teaching Assistant session. All students were given an answer key for all activities and homework. The results described show however that the effort is worthwhile. Not only were students more at ease in using the technology in the SOCR group, but they were more homogeneous in their performance and more satisfied than the students in the control group.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of this project by NSF DUE 0442992 and by an NIH Roadmap for Medical Research, NCBC U54 RR021813 grants. We thank Wei Sun and Xuelian Wei for their good suggestions and careful teaching assistance during the duration of the study.

REFERENCES

Aberson, C. L, Berger, D. L., Healy, M. R. and Romero, V. L. (2002). An interactive tutorial for teaching statistical power. *Journal of Statistics Education*, 10(3), http://www.amstat.org/publications/jse/v10n3/aberson.html.

Anderson-Cook, C. M. and Dorai-Raj, S. (2003). Making the concepts of power and sample size relevant and accessible to students in introductory statistics courses using applets. *Journal of Statistics Education*, 11(3), http://www.amstat.org/publications/jse/v11n3/anderson%2Dcook.html.

Cohen, S. and Chechile, R. A. (1997). Probability distributions, assessment and instructional software: Lessons learned from an evaluation of curricular software. In J. B. Garfield and I. Gal (Eds.), *The Assessment Challenge in Statistics Education*, (p. 253-262). Amsterdam: IOS Press and the International Statistical Institute.

Metz, K. E. (1997). Dimensions in the assessment of students' understanding and application of chance. In J.B. Garfield and I. Gal (Eds.), *The Assessment Challenge in Statistics Education*, (p. 223-238). Amsterdam: IOS Press and the International Statistical Institute.

Ross, S. (2006). *A First Course in Probability* (6th edition). Upper Saddle River, NJ: Prentice Hall.

West, R. W. and Ogden, R. T. (1998). Interactive demonstrations for statistics education on the world wide web. *Journal of Statistics Education*, 6(3), www.amstat.org/publications/jse/v6n3/west.html.