PROBABILITY

Collaborative Data Visualization Software

The data collection and analysis tool iSENSE helps algebra students collect, share, and explore their own experimental data while learning about the law of large numbers.

Melinda B. N. Willis, Sue Hay, Fred G. Martin, Michelle Scribner-MacLean, and Ivan Rudnicki

As mathematics teachers, we continually look for ways to make the learning of mathematics more active and engaging. Hands-on activities, in particular, have been demonstrated to improve student engagement and understanding in mathematics classes (Cruse 2013). Likewise, many scholars have emphasized the growing importance of giving students experience with the Internet and related technologies at school, which give us new ways to communicate and construct knowledge—for example, using search, collaboration tools, and discussion forums (Leu et al. 2004).

In the activity described here, we incorporated a free, Web-based data collection and analysis tool, iSENSE™, into a hands-on probability lesson for Algebra 2 students. Using iSENSE, students were able to collect, share, and explore their own experimental data, learning about the law of large numbers. (This theorem holds that as the number of observations increases, we can expect the experimental probability to converge to the theoretical probability.) Because they had collected the data themselves, the students participating in this activity felt a personal connection to the data and demonstrated a stronger grasp of the underlying math-

WITH


VLADIMIRS/THINKSTOCK

ematical principle. The activity addressed Common Core Standards for Mathematical Practices and mathematics learning standards in the area of high schools and probability (CCSSI 2010; see the sidebar **Common Core Mathematics Connections**).

The iSENSE system was developed by the Engaging Computing Group at the University of Massachusetts Lowell to enable students in mathematics, science, or other courses to easily collect, share, and explore multiple data sets online (Martin et al. 2010). The website, isenseproject.org, allows users to key in or upload both quantitative and categorical data and then to display their data in a variety of graphical formats, including a map, timeline plot, scatter plot, bar chart, pie chart, and histogram. Each visualization mode offers a number of configuration settings as well as the ability to automatically calculate means and medians, use logarithmic axis scales, and overlay lines of best fit. Most important for our use, the system can combine multiple data sets, contributed by different students, into a single graph. The teacher and students are able to explore data in real time and can achieve deeper class discussions than are typical with more traditional pencil-and-paper graphing exercises. Although not required, a classroom projector or

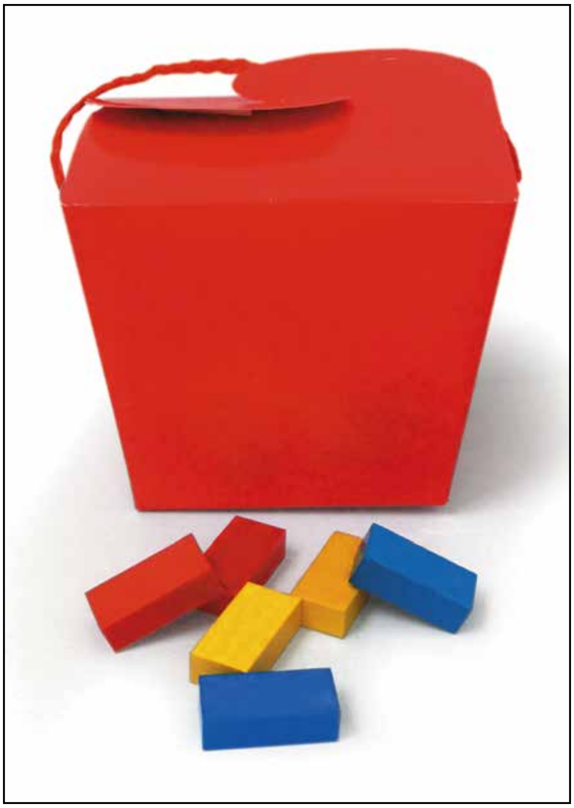

Fig. 1 The probability experiment uses materials that are readily available.

interactive whiteboard (such as a SMART Board™ or ēno® Board) greatly enhances iSENSE activities, allowing the teacher and students to analyze data as a group, iteratively configure and reconfigure data visualizations, and consider the effect of real-time data input.

We conducted three variants of a classic probability experiment using the iSENSE system. In each variant, students blindly drew colored blocks from a set of 12 blocks containing three different colors in varying proportions (e.g., 3 red blocks, 6 blue blocks, and 3 yellow blocks). In the first variant (Single Draw with Replacement), students drew a single block for each trial, recorded its color, and then replaced the drawn block before conducting the next trial. In the second variant (Multiple Draw with Replacement), students drew 2 blocks for each trial, recording the color of the first, replacing it, drawing a second block, and then recording the color of the second block. In the third variant (Multiple Draw without Replacement), students drew 2 blocks for each trial, recording the color of the first, not replacing it, then drawing a second, and recording its color as well. Student instructions for all three lessons are available on the iSENSE web site as part of the project called Block Probability.

SERGEY FESENKO/THINKSTOCK

This article describes the simplest variant of the experiment: Single Draw with Replacement. Before engaging in the activity, participating students had completed two days of lessons and homework on the fundamentals of theoretical probability. A review of theoretical versus experimental probability may be found on the website www.algebra-class.com (Hutchinson 2013).

DATA COLLECTION AND ENTRY

Students conducted the probability experiment in teams of two. Each team received a takeout food box containing 12 colored blocks, all identical to one another except for their colors (see **fig. 1**). In every set of 12 blocks, there were 3 red blocks, 6 blue blocks, and 3 yellow blocks. Therefore, the theoretical probability of drawing a red block was 3 in 12 (0.25, or 25%); the theoretical probability of drawing a blue block was 6 in 12 (0.50, or 50%); and the theoretical probability of drawing a yellow block was 3 in 12 (0.25, or 25%).

The students were told that each box contained 12 blocks but were not told how many blocks of each color were in the box. Over a series of 10 trials, students selected a single block from the box, recorded its color, and then placed the block back into the container. Students then used their individual data sets to make a prediction of how many blocks of each color were in the container, using the following equation:

> $(N/10) = x$ where N is the total number of a color selected; 10 is the number of trials conducted; and x is the resultant experimental probability

For example, if a group drew 2 red blocks in 10 trials, the equation yielded an experimental probability of 2/10 (0.20, or 20%), leading students to predict that the box of 12 blocks contained 0.20 • 12, or 2.4 red blocks. A student team drawing 4 red blocks in 10 trials would derive an experimental probability of 4/10 (0.40, or 40%) and make a corresponding prediction of 4.8 red blocks in the box. Because these predictions were calculated on the basis of only 10 observations, we encouraged individual teams to place only a small amount of confidence about the accuracy of this prediction.

After calculating their experimental probabilities independently, the student teams entered their results into the iSENSE website, which provides a simple interface for keying in numerical data

(see **fig. 2**). So that they could easily identify their team's data, each pair of students was instructed to enter their first names as the dataset name. Students found this process logical and simple, given their familiarity with technology in general.

Once the data were entered, students explored their data on the website, using different visualization tools. Students were encouraged to consider how different graphing formats affected their original analyses and inferences. During this open exploration time, we heard students discussing with one another which visualizations they preferred and why, a topic the class then considered as a group using a SMART Board.

WHOLE-CLASS ANALYSIS

From the SMART Board, we accessed the Blocks Probability project on iSENSE, where all the contributed data were available for analysis. The SMART Board enabled us to quickly and easily maneuver around the visualizations, adding or omitting data sets and configuring the graphs in front of the class. **Figure 3** is a screen shot of a sample of the data collected for the activity.

Each bar in the graph represents a number entered by a student team. The cluster of bars on the left shows the number of red blocks drawn by each team in 10 trials; the cluster in the center shows the number of blue blocks drawn by each team; and the cluster on the right shows the number of yellow blocks drawn by each team. Within a cluster, the color of each bar corresponds to the identity of the team that contributed the data, allowing the teacher to see which data set belongs to which student team. Mousing over a data bar brings up a pop-up dialog box, displaying its exact value as well as the identity of the contributing team. This feature adds to the classroom discussion by making it easy for each team to see and discuss its data. It is also helpful for identifying any erroneous data points, which can then be deleted or corrected.

SERGEY FESENKO/THINKSTOCK

To start the discussion, we initially chose data from only one group to demonstrate that the experimental probability resulting from just 10 trials was not likely to closely match the theoretical probability. The first bar in the cluster

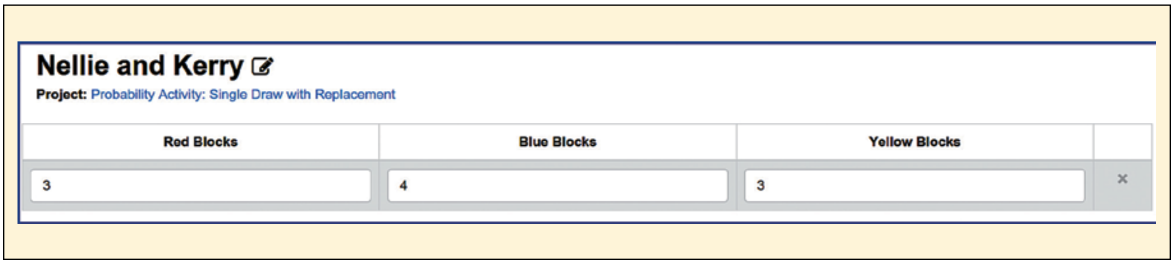

Fig. 2 The interface allows for data entry.

Fig. 3 Class data are aggregated on iSENSE.

Connections to the Common Core State Standards for Mathematics

The activity discussed here has connections to four Common Core mathematics content standards (High School: Statistics and Probability) and four mathematical practices standards (CCSSI 2010):

Standards

Interpreting Categorical and Quantitative Data

Summarize, represent, and interpret data on a single count or measurement variable

- Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets. (CCSS.MATH.CONTENT.HSS.ID.A.2)
- Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers). (CCSS.MATH.CONTENT.HSS.ID.A.3)

Making Inferences and Justifying Conclusions

Understand and evaluate random processes underlying statistical experiments

- Understand statistics as a process for making inferences about population parameters based on a random sample from that population. (CCSS.MATH.CONTENT.HSS.IC.A.1)
- Decide whether a specified model is consistent with results from a given data-generating process—for example, using simulation. If a model says a spinning coin falls heads up with probability 0.5, would a result of 5 tails in a row cause you to question the model? (CCSS.MATH.CONTENT.HSS.IC.A.2)

Mathematical Practices

The Probability with Collaborative Data Visualization Software activity engaged students in problem solving, reasoning, modeling, and use of software tools, connecting with these SMPs:

1. Make sense of problems and persevere in solving them. (CCSS.MATH.PRACTICE.MP1)
2. Construct viable arguments and critique the reasoning of others. (CCSS.MATH.PRACTICE.MP3)
3. Model with mathematics. (CCSS.MATH.PRACTICE.MP4)
4. Use appropriate tools strategically. (CCSS.MATH.PRACTICE.MP5)

on the left shows that there were 2 red draws out of 10 trials, yielding a probability of 0.20, or 20%, with a corresponding prediction that 2.4 of the 12 blocks in the box were red. We then selected two additional data sets for visualization. On the basis of this slightly larger sample (7 red blocks in 30 draws), the whole class redid its calculations and derived a combined experimental probability of 0.233, or 23.3%, closer to the theoretical probability of 25%. The ease with which the data sets were selected or deselected allowed meaningful analysis of the bar graph and how it changed according to the number of data sets included.

Students were eager to look at different combinations of data and wanted to operate the SMART Board themselves. They were given the chance to come to the front of the room and choose varying combinations of data to discuss. Each different combination led to a discussion about the experimental data and their relationship to the theoretical probability. The students in later classes were able to work with data from earlier groups. The final class observed 63 red blocks in 230 draws. The students calculated an experimental probability of 0.273, or 23.7%, corresponding to 3.28 red blocks in the box.

The combination of the iSENSE website and SMART Board technology enabled us and the students to easily enter the data, manipulate which data sets were being analyzed, correct input errors on the fly, combine or separate the individual data points, and quickly display our results. Students found that the hands-on activity and the analysis of many combinations of data enhanced their understanding of probability and, more specifically, their understanding of the effect of the number of samples drawn on the accuracy of the result. The students witnessed the law of large numbers in action as the experimental probability of the class data approached the theoretical probability.

SERGEY FESENKO/THINKSTOCK

USING iSENSE FOR OTHER ACTIVITIES

The iSENSE website is an open platform for the exchange of classroom activities involving data collection as well as the data collected in those activities. Although the experiment described here involved students from a single school, the website supports data sharing among multiple schools as well. (In theory, the more trials that are contributed to our project, the more closely the experimental probability will converge on the theoretical probability.) Any teacher who creates a free user account on the system may contribute data to our project, clone the activity for use in his or her class, or design a completely novel activity.

On the iSENSE website, all data are contributed in the context of "projects," which identify the

source of the data or how the data for the project should be collected. When creating a project, a user must define at least one data "field." Each data field must be assigned a specific type—time stamp, number, text, or location—as well as a unique name to identify the field. For any data field that is designed to store numbers, users also may specify a unit of measurement. For example, if students will be measuring the length of a group of objects, the type of the data field would be number, its name could be "length," and the unit of measurement could be inches. Our project had just three numerical fields: red blocks, blue blocks, and yellow blocks.

The types of data fields included in a project will determine the types of data visualizations available when users view the data contributed to that project. Most visualization tools require at least one number field to work. Text fields are used to store text values. Text values can be used to group, filter, and summarize number values. Location fields are used to store location information, such as latitude and longitude values, and address information. Data sets in projects with location fields can be viewed using the Map visualization tool. It is also possible to upload other resources, such as PDF or Word documents, photographs, and other media to an iSENSE project.

Step-by-step tutorials for creating projects on iSENSE, contributing data sets, and configuring data visualizations can be found on the website www.isenseproject.org/tutorials. The specific iSENSE project for this activity, including additional classroom support materials, is available at isenseproject.org/projects/336. Teachers interested in using iSENSE in their classrooms are encouraged to contact the authors for further information and advice.

ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under grant numbers IIS-1123972 and IIS-1123998.

REFERENCES

Charles, Randall I., Dan Kennedy, Basia Hall, Stuart J. Murphy, Grant Wiggins, Allan E. Bellman, Sadie C. Bragg, William Handlin Sr., Laurie E. Bass, and Art Johnson. 2012. *Algebra 2 Common Core*. Boston: Pearson.

Common Core State Standards Initiative (CCSSI). 2010. Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf

Cruse, Andra R. 2013. "Using Hands-On Learning Activities in High School Mathematics Classes to Impact Student Success." ProQuest Information and Learning. Dissertation Abstracts International 73(12-A)(E).

Hutchinson, Karin. 2013. http://www.algebra-class.com/theoretical-probability.html.

Leu, Donald J., Jr., Charles K. Kinzer, Julie L. Coiro, and Dana W. Cammack. 2004. *Toward a Theory of New Literacies Emerging from the Internet and Other Information and Communication Technologies. Theoretical Models and Processes of Reading.* 5th ed. International Reading Association. http://www.readingonline.org/newliteracies/leu

Martin, Fred, Sarah Kuhn, Michelle Scribner-MacLean, Christopher Corcoran, James Dalphond, John Fertitta, Michael McGuinness, Sam Christy, and Ivan Rudnicki. 2010. "iSENSE: A Web Environment and Hardware Platform for Data Sharing and Citizen Science." Paper presented at the AAAI 2010 Spring Symposium, Educational Robotics and Beyond: Design and Evaluation, Stanford, CA.





MELINDA B. N. WILLIS, melinda_willis@student.uml.edu, teaches K–14 mathematics and works with preservice mathematics and science teachers at schools and universities in New Hampshire and Massachusetts. **SUE HAY,** sue.hay@sau41.org, teaches mathematics and engineering courses and leads the FIRST robotics team at Hollis Brookline High School in Hollis, New Hampshire. **FRED G. MARTIN,** fredm@cs.uml.edu, professor of computer science at the University of Massachusetts Lowell, is director of the university's Engaging Computing Group and the principal investigator for the iSENSE project. **MICHELLE SCRIBNER-MACLEAN,** michelle_scribnermaclean@uml.edu, clinical associate professor of science and math education at University of Massachusetts Lowell, led the educational component of the iSENSE research project, bringing together the creativity of the developers and the realities of real-world application in the classroom. **IVAN RUDNICKI,** ivan@machinescience.org, is co-founder of Machine Science Inc., a nonprofit organization focusing on after-school enrichment programs for urban public school students.