Clifford Konold

Teaching Probability through Modeling Real Problems

In an attempt to reduce the growth of its population, China has instituted a policy that limits a family to one child. This policy has been particularly unpopular among rural Chinese, who have suggested revising the policy to limit families to one son. Suppose you were among those in the government considering the implications of adopting this proposal. Two questions you would no doubt ask yourself about the effects of instituting this "one son" policy are these:

1. What would be the average number of children in a family?
2. What would be the ratio of births of girls to births of boys?

The questions associated with this problem generate considerably more enthusiasm than problems about dice and coins that typify introductions to probability. This article describes a lesson that exemplifies an alternative approach to teaching introductory probability. In this approach, students learn to apply probability models to real-life situations and estimate probabilities through conducting simulations. (See NCTM [1981] for several articles on using simulations in teaching probability.) The particular activity described in this article has been used in high school and introductory college courses for which Macintosh laboratories and the simulation tool Prob Sim™ (1992) were available. However, it could be done using other software, or without computers, by having students model the problem by flipping coins and pooling the class's data.

An excerpt from Kristof (1990)

More in China Willingly Rear Just One Child

by Nicholas D. Kristof

DUANJIABA, China—Gong Daifang, a 36-year-old peasant, has made a decision that her ancestors probably would never have understood: her 10-year-old son is enough, and she will not have another child.

Even when local leaders in this brick-and-mud village in Sichuan province offered to allow her to have a second child, she decided that for practical economic reasons she would have only one, instead of seven like her own mother.

"My husband and I thought it over, but we believe that this way we can save money and lead a better life," Miss Gong explained.

China's population, the world's largest at 1.1 billion, is still increasing. But there are indications of a revolution in attitudes, with more and more Chinese couples falling in line with the nation's one-child policy as a matter of choice rather than compulsion. In view of the changing attitudes, some experts are predicting that China's population will actually decline after peaking in the early 21st century.

Coercion still underlies the one-child policy, and the rationing of the right to become pregnant remains a source of tension and bitterness in many parts of China. Many peasants grumble that the policy is not always carried out fairly, or should not be applied to them until they give birth to a son. Even Government officials acknowledge that some women are probably still forced to have abortions, and that many parents would like more children than they are allowed.

Fig. 1

A DESCRIPTION OF THE CLASS ACTIVITY

I begin this class with a discussion of a segment from an article published in the *New York Times* (Kristof 1990), which includes a description of the one-son policy **(fig. 1).** Students are asked individually to give estimates for the two values in the foregoing questions. Common responses for the average number of children are three or four children; a typical guess for the ratio of births of girls to births of boys is 2 to 1.

After students have discussed and argued for their various estimates, they simulate the one-son policy by drawing with replacement from a "sampling box" that has one B and one G. The number of elements drawn until a B occurs corresponds to the number of children in a simulated family (see the Prob Sim™ screen in **fig. 2**). Students summarize

Clifford Konold teaches at the University of Massachusetts, Amherst, MA 01003. He has been developing probability and data analysis curricula and software with funding from the National Science Foundation (Grant No. MDR-8954626). The opinions expressed are those of the author and not necessarily those of the Foundation. The author thanks students at Holyoke High School, Holyoke, Massachusetts, and their teacher, Al Gagnon, for allowing him into their classroom, and Ruma Falk, for helpful comments on an earlier draft.

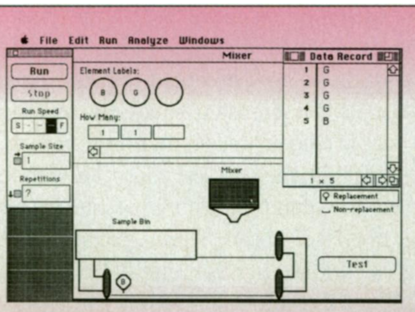

Fig. 2
A screen shot of Prob Sim™ being used to model the one-son problem. The "mixer" window shows one boy (B) and one girl (G) in the mixer, from which samples are drawn with replacement. The "data record" window shows the results of the last trial, which required five repetitions (children) before a B occurred. Each time the "run" button on the left is pressed, the computer keeps drawing until a B is selected. Because the length of the trial is unknown, a question mark (?) is displayed in the repetitions field on the right of the screen. The apparatus at the bottom of the mixer window animates the sampling process.

the results of 100 trials in a bar graph similar to the one shown in **figure 3.**

The students are asked to determine the total numbers of boys and girls and the average number of children in a family. The students have had some difficulty determining the total number of boys and girls from the bar graph. Some realize that since each family has exactly one boy, the number of boys is simply 100, but more thought is required to determine the number of girls from the heights of the various bars. The computation of the average proves even more troublesome. Some students become confused about the meaning of "average number of children in a family" in the context of the bar graph. Part of this confusion may have to do with the fact that the unit "family" is not visible as a discrete entity in the bar graph. In addition, data presented in the bar graph do not allow application of the standard add-and-divide algorithm, and many students have never knowingly calculated a weighted average. Whereas some students are eventually able to compute these values working only with the bar graph, others opt to refer to the original data record, which lists the various "family" sizes as they occurred during the simulation.

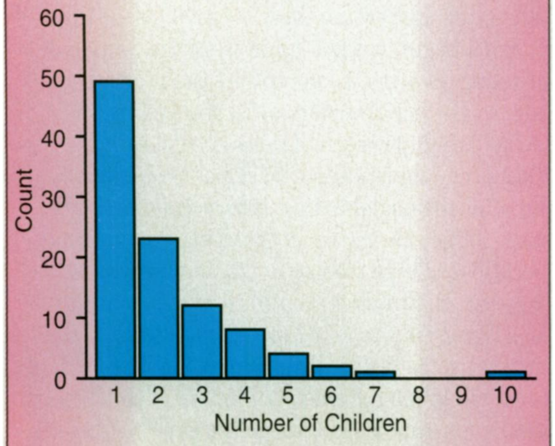

Fig. 3
Frequencies of number of children in a family in 100 simulated "families" under the one-son policy

On the basis of the shape of the bar graphs, students formulate the conjecture that the number of families having a given number of children decreases by a factor of 2 for each additional child. The suggestion that they think about the fraction of families that would have 1, 2, . . . , k children can help students verify the fact that the anticipated fraction halves as one moves from k to $k + 1$. They can use this information to predict what fraction of families would consist of, say, twelve children. The abler students will summarize the rule by the expression $1/2^k$.

What are the effects of the "one son" policy?

After doing these computations, students find that the total numbers of boys and girls are about the same and that the average number of children in a family is close to 2. When confronted with these surprising results, students are motivated to explore the problem more formally.

IMPORTANT FEATURES OF THE PROBLEM

Several features of this problem help to maintain students' interest and promote conceptual development. First, the results are counterintuitive. Finding a surprising result, students are more motivated than they otherwise would be to understand what is going on. They eagerly express opinions in class discussions.

Second, the problem offers multiple options for further analysis, both rudimentary and advanced. Usually, the first step in understanding the problem is for students to realize that their expectation of more girls than boys results from imagining large families consisting of one boy and many girls. They tend to overlook the fact that half of all families will have only one child—a boy. But students are capable of constructing more formal arguments. For example, two high school students independently formulated the following solution, perhaps inspired by their bar graphs. Among the first-born children, one would expect an equal number of boys and girls. The boys, of course, would all be the only child in the family. The families with first-born girls would all go on to have another child, of which approximately half would be boys and half girls. Therefore, if we looked at all second-born children, we would expect equal numbers of boys and girls. The same argument can be applied to third-born children, and so on. In each birth-order group, we expect an equal number of boys and girls. Therefore, we should expect the total number of boys and girls in the population to be equal.

Another argument involves a thought experiment (or, perhaps, a class activity). Imagine passing a coin along a line of people who, in turn, flip the coin once and pass it to the next person. A hypothetical result involving thirty people is shown at the top of **figure 4.** Everyone will agree that the expected numbers of Hs and Ts obtained in this process are equal.

Let T represent a female birth and H a male birth. All those who flipped a T come forward and line up in front of the next H that occurs down the line. Following this convention, the line of Hs and Ts would be restructured as shown at the bottom of **figure 4.** The people in each vertical line can be thought of as a family resulting from the one-son policy. The clusters are formed in the same way—by continuing on until the occurrence of an H. Yet we know from the way in which the data were generated that the outcomes H and T should be roughly equal in number. Some readers will realize on seeing this demonstration that they should have known the result all along—that the expected ratio of 1 to 1 follows from the fact that the outcomes of successive births are independent of one another. A similar problem, "Do men have more sisters than women?" was discussed by Falk (1982). Notice that in the example given, the sequence of thirty flips conveniently terminates with an H. This outcome is necessary if we insist that every family have a boy. If done as a class demonstration, this outcome can be accomplished by cutting the sequence off at the last H. However, equal numbers of heads and tails are expected whether the original sequence ends in an H or a T. In the situation we are modeling, the way to interpret the end of the segment that consists of n Ts (e.g., . . . HTTT) is as a family of n girls that, as of yet, has not had a boy. But according to the aforementioned birth-order argument, the existence of such families in the population does not affect the expected equal ratio of boys to girls.

Students discuss their predictions, then do the simulation

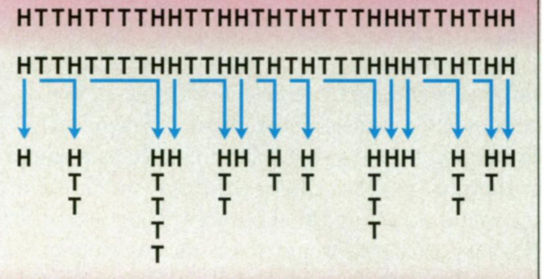

Fig. 4
Thirty flips of a coin (top) reorganized (bottom) to simulate "families" under the one-son policy

Once students have accepted that the expected number of boys and girls is equal, the expected family size is easy to derive. In n families with children born under the one-son policy, the number of boys must be equal to n. Therefore, the expected number of girls is also n. Thus, the expected total number of children in n families is $2n$, or an average of $2n/n = 2$ children per family. Note that by starting with the fact that the average family size is 2, we can show by the reverse reasoning that the expected numbers of boys and girls are equal.

Another important feature of the one-son problem is that it involves modeling a real situation. Students often balk when given the standard introductory problems: "What has this got to do with anything?" This comment is not to say that having students seriously consider the standard problems is unimportant. Indeed, given that coin-flipping is the prototypical chance event, good reasons can be cited to study it in some detail. But if we want to demonstrate the broad range of probability applications, then the situations we ask students to consider must become more complex than flipping coins, rolling dice, and blindly selecting socks from drawers.

A more important reason to have students model real situations is that they spontaneously raise various objections to modeling births under the one-son policy:

- What about twins?
- Doesn't a slightly higher chance exist for a male birth?
- Not all couples will keep trying until they have a boy.
- Who's going to make sure that a couple stops with one son?
- What about couples who can't have children?

Through addressing these types of objections, students begin to understand what the process of modeling involves. To determine the predictive value of the simulation results, students must decide if and how each of these considerations affects the real situation. In some instances they can alter the model to take into account an additional factor, for instance, adjusting the probabilities of B and G or placing a maximum value on k. They can also predict the direction of biases introduced by their simplified model; for example, the fact that some families will stop before the birth of a son will lower the average number of children per family, but will it also change the expected gender ratio of 1:1? Through this process of comparing a model to the target situation, students can come to realize that they can't avoid simplifying assumptions, but that the more aware they are of the limits of a particular model, the more informative the data from that model become.

REFERENCES

Falk, Ruma. "Do Men Have More Sisters Than Women?" *Teaching Statistics* 4 (1982):60–62.

Kristof, Nicholas D. "More in China Willingly Rear Just One Child." *New York Times,* May 1990, A1, A8.

National Council of Teachers of Mathematics. *Teaching Statistics and Probability.* Reston, Va.: The Council, 1981.

Konold, Clifford, and Craig Miller. Prob Sim™. Amherst, Mass.: 1992. Available from Scientific Reasoning Research Institute, Hasbrouck Laboratory, University of Massachusetts, Amherst, MA 01003.