Lesson 17: How do we build a simulation?

One-Boy Family Planning

Suppose that couples who wanted children were to continue having children until a boy is born. Assuming that each newborn child is equally likely to be a boy or a girl, would this behavior change the proportion of boys in the population? This question was posed in an article that appeared in *The American Statistician* (1994: 290–293), and many people answered the question incorrectly. We use simulation to estimate the long-run proportion of boys in the population if families were to continue to have children until they have a boy. This proportion is an estimate of the probability that a randomly selected child from this population is a boy. Note that every sibling group would have exactly one boy.

Digits
boys even 0,2,4,6,8
girls odd 1,3,5,7,9

There are seven steps to a simulation…

1. Identify the component to be repeated.
2. Explain how you will model the component's outcome.
3. Explain how you will combine the components to model a trial.
4. State clearly what the response variable is.
5. Run several trials.
6. Collect and summarize the results of all the trials.
7. State your conclusion.

Groupwork

	Components	Comments
Think	**Creates a successful simulation:** ○ randomizes the order of the envelopes ○ avoids giving any envelope out twice	
Show	**Conducts the simulation:** ○ describes the method clearly ○ shows the results of 20 trials, clearly labeled ○ defines the correct response variable	
Tell	**States a conclusion:** ○ establishes a reasonable decision rule ○ justifies the rule ○ does not confuse the model with the actual test to be conducted	

Components are scored as **E**ssentially correct, **P**artially correct, or **I**ncorrect

ESP

Your friend claims he "has ESP". Being properly skeptical, you decide to test his claim. Here is your plan.

You will get ten volunteers to sign their names on identical cards, and seal the cards in identical envelopes. You will then shuffle the pile of envelopes, and hand them to your friend. Using his alleged powers of extrasensory perception, he will distribute the envelopes back to the volunteers, trying to match each person with the one containing the proper signature.

Of course, it will be quite stunning if, when the ten volunteers open the envelopes, they all find their own signatures. If that happens you will certainly believe he really does have ESP.

But that's unlikely. Chances are he'll match some people with their signatures and miss others. You need to know how well an ordinary non-ESP-endowed person might do just by chance. Then you can decide how many matches your friend needs to make to convince you that he does have some mystical insight.

Before actually conducting this test then, you need to simulate it. You may use either your calculator or the random number table to determine how many matches you would consider to be "statistically significant". Write a report in which you clearly explain your procedure, show the results of at least 20 trials, and state your conclusion.

Kisses

Background: The paper "What Is the Probability of a Kiss? (It's Not What You Think)" (found in the online *Journal of Statistics Education* [2002]) posed the following question: What is the probability that a Hershey's Kiss will land on its base (as opposed to its side) if it is flipped onto a table? Unlike flipping a coin, there is no reason to believe that this probability is .5.

Working as a class, develop a plan for a simulation that would enable you to estimate this probability. Once you have an acceptable plan, carry out the simulation and use it to produce an estimate of the desired probability.

Do you think that a Hershey's Kiss is equally likely to land on its base or its side? Explain.

Simulation: Birthday Problem Suppose there are 30 people at a party. Do you think any two share the same birthday? Let's use the random-number table to simulate the birthdays of the 30 people at the party. Ignoring leap year, let's assume that the year has 365 days. Number the days, with 1 representing January 1, 2 representing January 2, and so forth, with 365 representing December 31. Draw a random sample of 30 days (with replacement). These days represent the birthdays of the people at the party. Were any two of the birthdays the same? Compare your results with those obtained by other students in the class. Would you expect the results to be the same or different?

A Crisis for European Sports Fans?

Background: The *New Scientist* (January 4, 2002) reported on a controversy surrounding the new Euro coins that have been introduced as a common currency across most of Europe. Each country mints its own coins, but these coins are accepted in any of the countries that have adopted the Euro as its currency.

A group in Poland claims that the Belgium-minted Euro does not have an equal chance of landing heads or tails. This claim was based on 250 tosses of the Belgium Euro, of which 140 (56%) came up heads. Should this be cause for alarm for European sports fans, who know that "important" decisions are made by the flip of a coin?

In this activity, we will investigate whether this difference should be cause for alarm by examining whether observing 140 heads out of 250 tosses is an unusual outcome if the coin is fair.

1. For this first step, you can either (a) flip a U.S. penny 250 times, keeping a tally of the number of heads and tails observed (this won't take as long as you think) or (b) simulate 250 coins tosses by using your calculator or a statistics software package to generate random numbers (if you choose this option, give a brief description of how you carried out the simulation).
2. For your sequence of 250 tosses, calculate the proportion of heads observed.
3. Form a data set that consists of the values for proportion of heads observed in 250 tosses of a fair coin for the entire class. Summarize this data set by constructing a graphical display.
4. Working with a partner, write a paragraph explaining why European sports fans should or should not be worried by the results of the Polish experiment. Your explanation should be based on the observed proportion of heads from the Polish experiment and the graphical display constructed in Step 3.

A "Hot Hand" in Basketball

Background: Consider a mediocre basketball player who has consistently made only 50% of his free throws over several seasons. If we were to examine his free throw record over the last 50 free throw attempts, is it likely that we would see a "streak" of 5 in a row where he is successful in making the free throw? In this activity, we will investigate this question. We will assume that the outcomes of successive free throw attempts are independent and that the probability that the player is successful on any particular attempt is .5.

1. Begin by simulating a sequence of 50 free throws for this player. Because this player has a probability of success = .5 for each attempt and the attempts are independent, we can model a free throw by tossing a coin. Using heads to represent a successful free throw and tails to represent a missed free throw, simulate 50 free throws by tossing a coin 50 times, recording the outcome of each toss.
2. For your sequence of 50 tosses, identify the longest streak by looking for the longest string of heads in your sequence. Determine the length of this longest streak.
3. Combine your longest streak value with those from the rest of the class, and construct a histogram or dotplot of these longest streak values.
4. Based on the graph from Step 3, does it appear likely that a player of this skill level would have a streak of 5 or more successes sometime during a sequence of 50 free throw attempts? Justify your answer based on the graph from Step 3.
5. Use the combined class data to estimate the probability that a player of this skill level has a streak of at least 5 somewhere in a sequence of 50 free throw attempts.
6. Using basic probability rules, we can calculate that the probability that a player of this skill level is successful on the next 5 free throw attempts is

$$P(SSSSS) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^5 = .031$$

which is relatively small. At first, this value might seem inconsistent with your answer in Step 5, but the estimated probability from Step 5 and the computed probability of .031 are really considering different situations. Explain why it is plausible that both probabilities are correct.

7. Do you think that the assumption that the outcomes of successive free throws are independent is reasonable? Explain. (This is a hotly debated topic among both sports fans and statisticians!)