USING SIMULATION, SPORTS, AND THE WWW TO HELP STUDENTS EXPERIENCE EXPERIMENTAL DESIGN

Robin H. Lock, St. Lawrence University, USA

Given an assumption that students can best learn and appreciate concepts involved with creating and analyzing experimental designs by actually performing their own statistical experiments, the question facing many instructors is, "How can we best provide interesting situations which allow students to practice the design techniques that are discussed in class"? We describe several settings that encourage students to investigate a clearly delineated problem, design an experiment (or series of experiments) to collect data to address the problem, and then analyze that data to move towards some resolution to the problem. Situations include a simple task to "fill" a coin with water, efforts to improve one's golf game, and more elaborate simulations in which multifactor experiments can be performed over the World Wide Web.

INTRODUCTION

What are the characteristics of an effective experimental situation for getting students to experience the trials and tribulations of statistical design? Obviously, much depends on the level of the course, concepts to be demonstrated, and the backgrounds, abilities and interests of the students. In what follows, we offer a few general guidelines for designing student experiences and we provide examples of three specific experimental situations that can be adapted to a variety of course levels.

SOME DESIRABLE CHARACTERISTICS

1. *A clear goal.* The problem should have an easily measured response variable with a well-understood objective -- often to minimize or maximize the response. Students enjoy the challenge of trying to find optimal treatment conditions. There should be a good amount of variability in the responses and at least some dependence with one or more of the explanatory factors.
2. *Multiple factors*. Students should be able to easily manipulate and control one or more explanatory factors, while other confounding or nuisance factors may be present. Additional factors and interactions between factors highlight the need to do a well-designed experiment as opposed to collecting happenstance data.
3. *Results which are surprising, non-intuitive, or unanticipated.* To maintain student interest in investigating the problem, the findings should show them something they would not have predicted in advance.

4. *Practical setup*. Students should be able to run various treatments without a great deal of time, effort, or expense. While it is important for students to realize that costs often seriously inhibit data collection -- thus necessitating a careful design -- we would rather have them spend time and effort on designing the experiment and interpreting the results, rather than collecting the data.
5. *Realism*. While it may be difficult to construct a classroom or lab experience which addresses questions that are important or relevant to each student's individual academic or personal interests, the situation should allow students to envision how they might adapt the techniques illustrated by the design exercise to problems in their own field.

EXPERIMENT #1: FILLING A COIN.

We use this experiment as a lab exercise to provide an initial exposure to ideas of statistical experimentation during the first month of a basic introductory course in applied statistics. The motivating question is stated very simply, "How many drops of water can we fit onto the face of a coin before the water overflows"? Materials include a medicine dropper, a glass of water, a coin, and sufficient paper towels to soak up the spills. A quick demonstration in front of the class shows that the answer is likely to be quite a bit larger than most students would expect. Surface tension allows the thickness of the "pile" of water on top of the coin to exceed the thickness of the coin and 20 or more drops are often required to cause an overflow for a standard U.S. penny.

What factors might affect the number of drops required? Students can easily brainstorm to come up with a number of candidates including:

⇒ the size of the opening in the medicine dropper
⇒ the denomination of the coin
⇒ the age of the coin
⇒ the condition of the coin
⇒ the temperature of the water or coin
⇒ the angle at which the dropper is held
⇒ the height from which the drops are released
⇒ the steadiness of the dropper operator

Since this experiment is intended as only a brief introduction for beginning students, they work in groups of two or three to design an experiment to investigate just a single factor of their choice. For example, one group might decide to study the effect of salinity by comparing plain tap water with water that has been salted, while another group might compare the carrying capacity of heads vs. tails. Although they are only studying a single factor, they are encouraged to discuss other factors in their lab reports and indicate how their experimental procedures either controlled possible extraneous factors or used randomization to minimize their effects. For example, a group might choose to have the same student do the dropping at each trial to minimize operator variation, they might decide to randomly allocate twelve coins into two treatment groups of size six, or they might choose a matched pairs design to account for wear in the head/tail comparison.

At this early stage of the course, the students have not yet seen formal techniques of inference so the analysis of their data is generally limited to summary statistics and graphical displays. However, we can motivate general concepts such as what might be a statistically significant effect while delaying more detailed study until later in the course when we may refer back to the original experiment. Also, by actually performing the experiment, students are likely to encounter more subtle design considerations such as the effect of learning on later trials if they gain experience with keeping drops on the coins or perhaps a fatigue factor if they get less careful in later runs.

EXPERIMENT #2: CHIPPING IN GOLF

This experiment has been derived from a project undertaken by a group of engineering managers who were attending a summer institute that included a five-week course in experimental design. Several of these students were avid golfers who decided that they could maximize their enjoyment of the required project by designing an experiment that could be performed on the local golf course. While the topic may not carry the same appeal for all students, the issues are straightforward and even a non-golfer can participate fully.

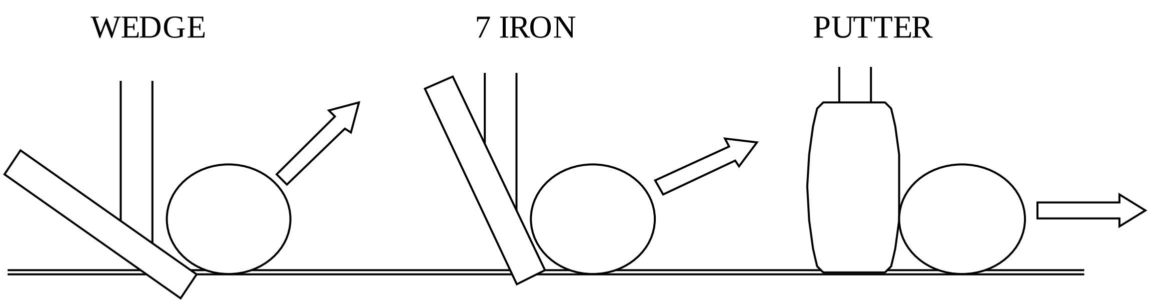
One of the goals in golf is to hit the ball onto the “putting surface” or “green” -- the area of closely mowed grass that surrounds the hole. Unfortunately, the inherent variation of golf shots causes even the very best players to miss this target, with the ball stopping short of the green, bouncing over it, or missing to one side. When this occurs, the player will often need to play what is known as a “chip” shot, requiring much less force than a full swing to put the ball on the green and (hopefully) stop it near the hole.

Among golfers there is considerable debate about the proper way to attempt this shot. Some recommend flipping the ball high in the air so that it flies most of the way to the hole, then lands softly and comes quickly to a stop. Others advocate a “chip and run” which uses a much lower trajectory and relies on the ball rolling for a longer distance along the smooth surface of the green. Occasionally, one might even use a flat putter to roll the ball the entire way. Even experienced golfers will vary the type of shot they use, depending on the conditions between the ball and the hole. Since the swing is fairly gentle and different clubs are designed to produce the various degrees of loft, even a novice golfer with just miniature golf experience can make reasonable attempts at these shots with just a bit of practice.

The response variable of interest is the *distance* that the ball stops from the hole. Obviously we would like to determine conditions which help us to minimize that distance. Once again, students will readily suggest potential factors such as:

- ⇒ the distance from the original ball position to the hole
- ⇒ the direction of slope (uphill, downhill, flat)
- ⇒ the proportion of projected path which includes the green (it's more difficult to get the ball to roll reliably if it hits before the green)
- ⇒ the club used to hit the ball (ex. lofted wedge, lower trajectory 7-iron, flat putter)
- ⇒ the thickness of grass at the ball (fairway, light rough, heavy rough)

An experiment that controls and manipulates some or all of these factors can easily be designed and carried out around the practice putting green of a local golf course (field trip!). Students can use a coin to move the "target" to various locations on the green and a tape measure to determine the response value for each shot. Because of the obvious potential for learning effects, they need to take special care to randomize the order of presentation of treatments. Students with some experience in playing golf readily



recognize the importance of gaining knowledge about the effects and interactions of the various factors to help make the best decision about the type of shot to attempt in a specific situation.

Even if there is not a convenient local golf course, your students have never swung a golf club, or your institution (like mine) is in a climate which finds the golf course more suitable for cross-country skiing during most of the semester, you could still try this experiment in a lab setting using one of the popular computer golf simulators such as Access Software's *Links* or Microsoft's *Golf*. Such programs frequently have a "practice" mode where you can set up specific situations, try out various clubs and swing styles, and see the results (including automatically computed distance to the hole) immediately on the screen after each shot. Of course, this becomes one step further removed from reality since you would now be attempting to model the behavior of the computer simulation, but some students might actually be more interested in seeing how statistical techniques could be applied to improve their computer game playing skills.

EXPERIMENT #3: MOTIVATION FOR PERFORMING ROUTINE TASKS

We can take the idea of simulation a bit further by imagining a computer program that controls all aspects of the design situation through simulation. Several interesting examples of such programs were written for the Michigan Experimental Simulation System (MESS) by Stout, Rajecki, Halstead-Nussloch, Main and others at the University of Michigan. A Web-based interface using cgi scripts has been developed for these programs by McClelland and Roberts at the University of Colorado's Computer Laboratory for Instruction in Psychological Research (CLIPR). Topics include imprinting in chickens, diagnosis of schizophrenia, the effects of motivation on performing routine tasks, and a two-stage experiment in social facilitation. They may be freely accessed at *http://samiam.colorado.edu/~mcclella/expersim/expersim.html*.

The table below gives a flavor of the setup for running the motivational factors simulation. Students click in boxes to specify their designs. Data (including codings for all factors) are automatically generated by the simulation program and displayed in a new screen from which students can save or cut and paste into a statistics package. A separate webpage describes the nature of the experiment (the response variable is the number of correct answers to arithmetic problems) and links provide definitions for key terms.

Motivational Factors in Routine Performance

Observations per condition	Size of testing subgroups	Length of testing task
________ *between 1-100 subjects*	❑ 1 (Individual testing) ❑ 5 people per session ❑ 25 people per session ❑ 100 people per session	❑ 10 minutes ❑ 30 minutes ❑ 50 minutes ❑ 60 minutes
Difficulty of task	Presentation of instructions	Control variables

❑ Simple ❑ Moderate ❑ Difficult	❑ Simple ❑ Important for science ❑ Unimportant	❑ Need for achievement ❑ Need for affiliation ❑ Fear of failure
Need for affiliation	Fear of failure	Need for achievement
❑ High ❑ Medium ❑ Low ❑ Evenly distributed	❑ High ❑ Medium ❑ Low ❑ Evenly distributed	❑ High ❑ Medium ❑ Low ❑ Evenly distributed
Year in school	Sex of experimenter	Sex of subjects
❑ Freshperson ❑ Sophomore ❑ Junior ❑ Senior ❑ Random selection	❑ Male ❑ Female	❑ Male ❑ Female ❑ Mixed (50-50) ❑ Random selection

Based on *http://samiam.colorado.edu/~mcclella/expersim/motivational.html*

CONCLUSION

The three experiments described here satisfy our criteria to varying degrees. Experiments #1 and #2 are certainly real, although one might question whether the topics of investigation are very important. Experiment #3 is the most *realistic* example of how experimental design can be applied in a research setting, but the data are being generated by a machine -- not real people. Each experiment allows for consideration of multiple factors and gives students easy control over some factors. Most students are not able to guess the relationships and effects in advance, so each experiment should show them how statistical design can help to gain new understanding about a physical situation. In each case rich data can be obtained with minimal expense. This is especially true of the online simulation and we would like to encourage further development of that sort of resource.