USING ACTIVITIES IN STAT 101

Jeffrey A. Witmer, Oberlin College, USA

The use of activities in introductory statistics courses is becoming more popular. These activities can take on many forms, ranging from activities that are done by the entire class as one group and that are led by the instructor to small-group work, to individual work that can be done out of class. In this paper we consider examples of different ways that activities can be used.

INTRODUCTION

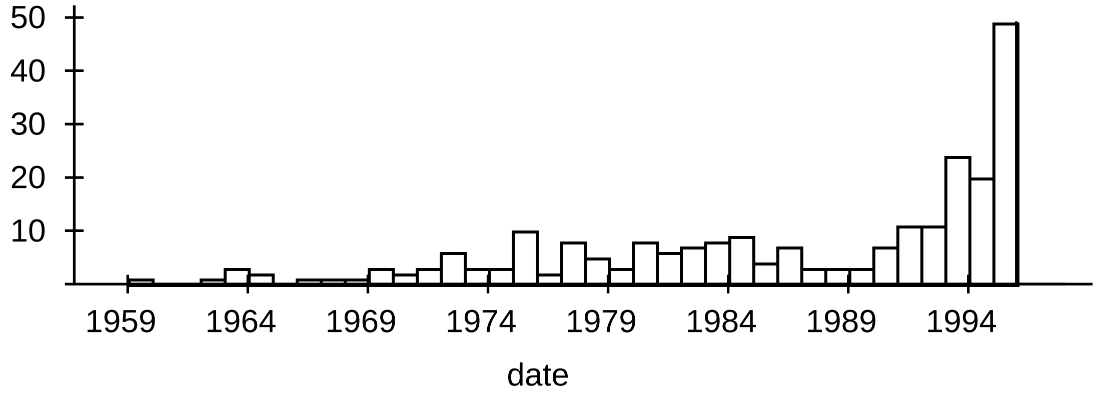
Statistics instruction is evolving in a number of ways. Some teachers are using student-directed projects, some are presenting less probability, some are making greater use of technology either in class, on assignments, or both, and some are making more use of real data. This paper will focus on a movement that runs parallel to these changes: the use of more activities in the classroom. Thus, the purpose of this paper is to consider the variety of ways in which activities can be used and to give some examples.

Experienced teachers have long recognized that for many students hearing a lecture, or reading a textbook, is less enlightening than engaging in an activity. The Chinese proverb "I hear and I forget, I see and I remember, I do and I understand" holds true for the study of statistics. Thus, the use of activities has been around for a long time (see Jowett and Davies, 1960). In recent years the interest in using activities as part of an introductory statistics course has grown. A number of projects have produced resources that teachers can use either in the classroom or in a lab room, with or without a computer. For example, see Rossman (1996), Scheaffer, et al., (1996), and Spurrier, et al. (1995).

Activities can be used in different formats. Some activities can be done by the entire class, whether the class is large or small. Other activities lend themselves to group work that can be done during class or as homework. There are some activities that are meant to be used within a lecture, with the teacher playing the leading role. Examples of these kinds of activities are discussed in the following sections.

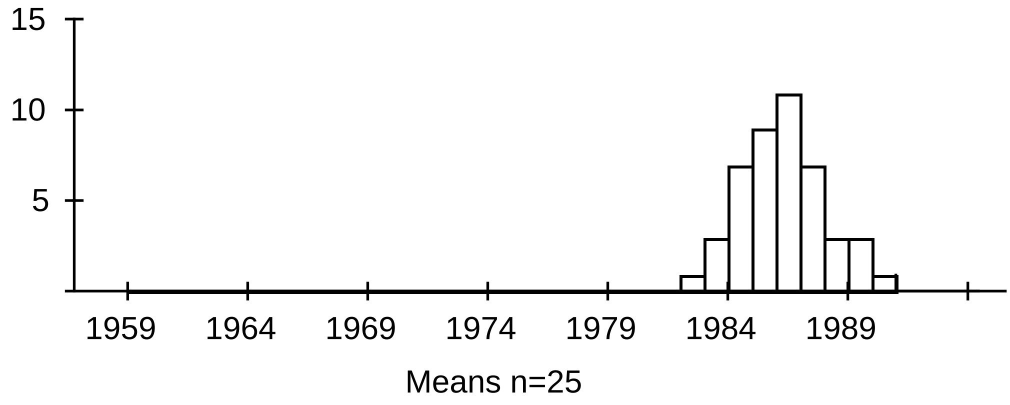
ACTIVITIES FOR THE ENTIRE CLASS

One of the activities in the book *Activity-Based Statistics* (Scheaffer, et al., 1996) is called "Cents and the Central Limit Theorem." This is a class exercise, which was developed by Ann Watkins, in which students bring collections of pennies to class and study the distribution of their ages. The activity starts by having the students form a histogram on the floor made of the pennies they have collected in the previous few days. The students place their pennies above a number line that spans the dates in which the pennies were minted. Figure 1 shows a typical distribution, which is heavily skewed.



The mean and standard deviation of this population are calculated by the teacher. The teacher then mixes the pennies together into a big pile and invites the students to take random samples of size 5 from the pile. Each student calculates the average date of her or his 5 pennies and places a nickel on the number line at the appropriate place. When the entire class does this, a histogram is formed that represents the sampling distribution of the mean when n = 5. The mean and standard deviation of this distribution are calculated by the teacher. The pennies are then mixed together and the process of random sampling is repeated, but using samples of size 10. A histogram of the means of n = 10 penny dates is formed on the number line using dimes. Finally, the students take samples of size 25 and build a histogram using quarters.

The class makes three observations at this time. (1) The means of the distributions when n = 5, 10, and 25 are all nearly equal to each other and to the mean of the original distribution. (2) The standard deviations of the distributions get smaller as n increases (and the square root of n gives the rate of decrease). (3) The histograms become bell-shaped when n gets large. Figure 2 shows this third point, which many students find surprising. It is remarkable that something so skewed can generate something bell-shaped -- a lesson the students are not likely to forget.



Another example of a whole-class activity is studying the sampling distribution of the slope in a regression problem. This requires every student having a computer or a graphing calculator to use during class time, so it might need to be done in a computer lab. Teachers whose students all have, for example, a TI-83 graphing calculator can do this activity in a regular classroom. The activity proceeds as follows. The teacher chooses a fixed regression model, such as Y = 10 + 2X + error, and chooses values for X, say 1, 2, ..., 12. Each student enters the values for X as one list and the values 10 + 2X as a second list in his or her calculator. Each student then generates a random error, from a normal population with mean 0 and standard deviation 4, say, for each of the 12 observations. The random errors are added to the values of 10 + 2X to get a new list, called Y. A regression of Y on X yields a fitted regression slope. The students call out the slopes they found and the teacher creates a dotplot of the values. It is quite easy to generate new random samples of errors and new sample slopes, so that each student can generate several sample slopes in just a few minutes time. Thus, the class dotplot can quickly have dozens of points in it. The fact that this dotplot takes on a bell shape and is centered at 2 is quickly apparent. One can also calculate the standard error of the slope and compare it to the standard deviation of the points in the dotplot.

INSTRUCTOR-LED ACTIVITIES

Some activities are orchestrated by the instructor. A quick activity that can be done by the entire class is to make a “Living Boxplot” using students. To do this, the class chooses a quantitative variable of interest to them that is easily measured on each student. The teacher marks off an imaginary number line on the floor at the front of the room and invites the students to stand at their place along the number line. The class then determines the median and the quartiles, inviting those students to step forward. After checking for outliers, the minimum and maximum students step forward to complete the shape of the boxplot. Such a living example of a boxplot helps the students see that the median and quartiles divide the number of points into quarters, not the number line itself. If the class is small, then every student can be part of the boxplot; for a large class one would only use a sample of the students.

"Introduction to Hypothesis Testing" is the name of another activity in *Activity-Based Statistics*. As is true of many of the activities in that book, versions of this activity have been used by many statistics teachers for many years. One version of the activity is as follows. The teacher tells the class that it is time to review randomization in a paired experimental design. Students are put into pairs according to where they are sitting in the room. The teacher takes out a deck of cards and says that if a randomly drawn card is red, then the first student in the pair will be assigned to the treatment group and the second student will be assigned to the control group. If the card is black, then those assignments will be reversed.

The teacher shuffles the deck and draws a random card, which is black. After replacing the card in the deck and shuffling again, another black card is drawn. This is repeated, with black cards drawn each time, until the class starts to laugh, which usually requires about 5 consecutive black cards. The teacher then asks the class "Why are you laughing?" and is told that it seems odd to get a black card every time. "But isn't this possible, even if the deck is a fair deck?" the teacher asks. "Yes," the students reply, "but it isn't likely." The teacher then explains that the class has demonstrated the logic of a hypothesis test -- reject a claim (perhaps by laughing) if the data are not consistent with what the claim predicts -- and that they have conducted their first hypothesis test. The p-value can be found at this point with a simple binomial calculation. The teacher *might* then choose to reveal how it is that every card drawn was black. One way to guarantee this result is to take to identical decks and switch the red cards in one deck for the black cards in the other deck.

INDIVIDUAL ACTIVITIES

Some activities can be done individually, perhaps as homework. An example of this kind of activity is "Streaky Behavior," in which the student is given a handout that contains two sequences of H's and T's, representing 50 coin tosses. (In the *Activity-Based Statistics* book the sequences are of length 200; I am describing a modified version of the activity.) One of the sequences is real but the other is made up. The made up sequence does not contain the kind of long runs of heads and tails that real coin tossing generates. Most students can't tell the sequences apart, so they are invited to toss a coin 50 times and record the results. A Stirling Recording Sheet, which gives a graphical history of 50 Bernoulli trial outcomes, is a handy place to keep a record of the tosses. After making 50 tosses, students generally see that long runs happen more often than intuition suggests; this helps then identify the real sequence.

As an extension of this activity, the teacher can get the class together and pose the following question. Suppose two students play a game of coin tossing, with the winner being the one with the most heads in 50 tosses. If we keep track of which player is leading the game as it is played, what is the chance that the lead will change hands at least once during the 50 trials? This is a hard problem to solve analytically, but lends itself to simulation. If Stirling Recording Sheets were used, then each student has a record of his or her coin-toss history. Pairing students together and having them hold their Recording Sheets up to a light gives a quick way to check on the sample percentage of paths on the sheets that cross. This sample percentage gives an estimate of the probability in question.

ACTIVITIES FOR GROUPS

"Matching Graphs to Variables" is an example of an activity in which students benefit from working in small groups. A set of 5 quantitative variables is presented, along with a set of 5 histograms and boxplots. The students are challenged to match the variables with their graphs. This requires thinking carefully about the features in a graph that one would expect for a given variable, such as bimodality or skewness.

MUST WE CHOOSE BETWEEN LECTURES AND ACTIVITIES?

Many teachers of introductory statistics organize their courses around a series of lectures in which they present concepts, formulas, and examples. This can be a highly effective way to convey information and, when done with skill, can be very impressive. Some teachers want to eliminate lectures as much as they possibly can. Others want to retain a lecture format, but use activities within their courses. The examples discussed above show various ways that activities can be used to augment the traditional lecture format. Any or all of these can be incorporated into an introductory course gradually, by evolution rather than revolution.

Some teachers have small classes but others must deal with very large classes. Some have access to computer labs but others don't have this luxury. Whatever the teaching environment and style of the instructor, it is important to bear in mind that when using activities the focus should be on the statistics concepts that are being taught. Students can easily lose sight of what they are doing and why they are doing it. This gives the teacher plenty to do, even if she or he chooses to lecture less.

REFERENCES

Scheaffer, M., Gnanadesikan, M., Watkins, A., and Witmer, J. (1996). *Activity-Based Statistics, Instructor Resources* . New York: Springer-Verlag.

Jowett, G. and Davies, H. (1960). Practical Experimentation as a Teaching Method in Statistics. *Journal of the Royal Statistical Society*, Ser. A, 123:10-36.

Rossman, A. (1996). *Workshop Statistics: Discovery with Data*. New York: Springer-Verlag.

Spurrier, J., Edwards, D., and Thombs, L. (1995). *Elementary Statistics Laboratory Manual*. Belmont, CA: Wadsworth.