

WHAT TO TEACH BEFORE INFERENCE: BUILDING THE BONES FROM PAPER, SOFTWARE, AND STORIES

Timothy E. Erickson and William F. Finzer, Key Curriculum Press, USA

In developing curriculum materials for the DataSpace project (supported by the US National Science Foundation awards III-9400091 and DMI-9660827), we are concerned that students build strong statistical “bones”—a robust framework of attitudes, skills, and concepts. We believe this goal is best served by a combination of (1) paper—well-chosen off-line activities; (2) software—flexible computer tools, especially for graphing, and (3) stories—many opportunities for students to make conjectures about data.

Many students find statistics mystifying; and many who pass the course don't understand what went on. This may be because of our zeal in getting to what we see as the real meat of statistics (doing inference well, for example) and giving short shrift to what we might call the “bones” of statistics. What are these bones? What should we do before formal inference so students will understand it? In the DataSpace project, we have been designing materials that try to address this issue using three interwoven threads:

- *Paper*—students make concrete manipulations using physical objects (often but not always slips of paper) to represent data.
- *Software*—students use the computer to increase the sizes of data sets and to perform and re-perform their analyses in a more abstract (but still informal) environment.
- *Stories*—students explain and describe the data in many ways, sometimes making literal “stories,” and sometimes making conjectures about the data.

Making paper manipulatives and using software are probably familiar to the reader. But what do we mean by *stories*? Rather than a definition, an example is in order—of a series of activities directed at developing the spine of a student's statistical skeleton: understanding distributions. This first activity uses paper and the simplest kind of story.

TURNING NUMBERS INTO PEOPLE

Students receive a sheet which, at first glance, appears to be an indecipherable grid of numbers. They quickly learn that each line represents an individual's data from the 1990 US Census. Furthermore, the lines are grouped together into households and there is a key that helps them decode some of the columns of numbers into real data. Here is an example household of three people (there is a “+” every ten columns, and these lines are about half the original length):

```

-----+-----+-----+-----+-----+-----+-----+
P0089307000001600001300000000101000000000481006000113022050100000002000
P008930701100236000260000000211020000000481036000314335999100000002000
P008930702100208400420000122100141000000481006000204335999100000002000
-----+-----+-----+-----+-----+-----+-----+

```

Column 11 is Sex (0=M, 1=F); Columns 12–14 are Race (001=White, 002=Black); Columns 15–16 are Age. This “sample” key gives you a taste—but you can learn a lot more. For example, yes, they are married. Their daughter is in public school. He has some college but no degree; she finished her Bachelor’s. And so forth.

The assignment is to prepare a story about the household you have chosen. Who are they? What are they like? These stories run from the pedantic to the racy, and let the teacher ask some important questions in an engaging context, for example: How sure are you about your story? What parts of your story are data and what are inferences? Where did you get the ideas that help you make your inferences?

Students and teachers in this exercise report that they find the previously impenetrable sea of numbers gradually resolves itself into a collection of real people.

HOUSEHOLDS OF CARDS

Decoding all those numbers is too hard to do for very long. We now give the students computer-printed cards of households in which the numbers are all converted into text. Each pair of students receives several dozen households, and, looking at the cards, tries to come up with conjectures—statements they believe may be true—about the set of data. Having done that, they make a display (often a two-by-two table populated by the cards themselves) that shows evidence whether the conjecture is true or false. A typical conjecture is, “Rich people live in big houses.” In this activity, the students discover:

- They may have to define their terms and restate their conjectures (“Rich people live in big houses” may become “a higher percentage of people with incomes over \$40,000 live in houses with five or more rooms”);
- That if individual cases do not support or refute a conjecture, the preponderance of cases still may;

Student-made conjectures are important: they promote interest in and ownership of the data. These conjectures are *stories* as well—proto-models that explain the data. In contrast to the previous activity,

0017528	
Detached one-family house; two vehicles 5 rooms with two bedrooms owned with mortgage or loan	
0017528	householder
Married (Mexican) Latino M age 37. Speaks Spanish at home. Postsecondary teacher/subject not specified earning \$55000. Unearned income: \$300. Master's degree	
0017528	spouse
Married Japanese F age 43. Editor/reporter earning \$11000. Bachelor's degree	
0017528	child
Japanese Latina F age 11. 5th, 6th, 7th, or 8th grade (in public school)	

these stories—the conjectures—help to summarize all of the data, not just bring one case to life. Here, “stories” are more like “newspaper stories” than biographical narratives.

The cards can introduce distributions of continuous variables. One activity that has been especially helpful is to have the students cut out the individuals and make a histogram out of the cards themselves—for example, a graph of the distribution of ages by decade.

Now we ask what features students see in the graph. They notice that there are fewer really old people than in the other bins. We ask for conjectures—for stories—to explain that feature of the distribution. Old people die, of course. Here, we reemphasize the difference between data and conjecture, and also insist on *multiple* conjectures. When pressed, students come up with more possible explanations: old people move away to other communities where more old people live; there were fewer people born eighty years ago, so of course their population isn’t as great; and so forth.

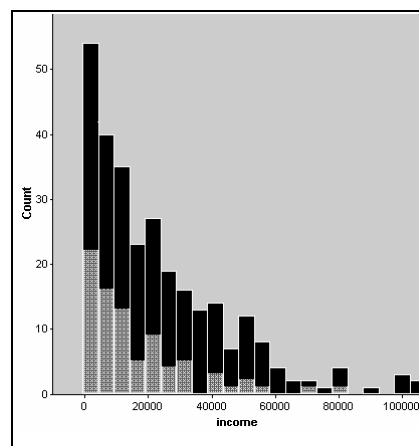
Though we have not yet (February 1998) done an experiment to test *our* conjecture, we believe that this lesson is very effective and that the combination of the paper and the stories makes this lesson work. The people are described individually on slips of paper, so the students think of the cases in the data set as actual people. More importantly, they think of the histogram bins or table cells as sets of people rather than as monolithic bars or aggregate numbers. Thinking about real people helps students bring outside knowledge (e.g., old people move away) to bear as they make up stories about the distribution. The result is that students have an interest and a purpose in characterizing the distribution of ages. Notice how different this is from a traditional introduction to distributions. We do not introduce terms like “symmetrical” or “skewed” at this time. Experiences with slips of paper support students later when symmetry and other properties of distributions become relevant.

WITH THE STUDENTS AT THE COMPUTER

At last, we let students use software in order to make graphs and investigate conjectures more easily. Having done the “paper” exercise, students seem to have a more secure understanding of what’s going on. We are more likely to get a good answer when we point to a (paper) card of an individual and a graph on the computer screen, and ask, “Where is this person in that graph?” Furthermore, the paper experience with distributions will, we hope, transfer to distributions the student never constructed offline.

For example, suppose we look at the income distribution of a Berkeley, California data set. If we plot the incomes of people between the ages of 20 and 70, we see this:

In this picture, we have 294 cases—too many to manipulate quickly on paper. We’ve colored the Whites in the sample darker (the tops of each bar). The rest, the minorities—mostly African-American—are colored gray. What story does this distribution tell us? One thing is that the *proportion* of the races in the bars gets whiter at higher incomes.



In the media, we tend to see only measures of center, e.g., Whites have a median income of \$21,000 while Blacks earn \$19,000. But a distribution tells us a richer story. Looking at this graph for a while, a workshop teacher commented that “there are still more poor Whites than poor Blacks”—a profound observation we would be unlikely to make without the help of software. Paper is the tool for introducing distributions; software is for exploring them.

ONWARD TO INFERENCE

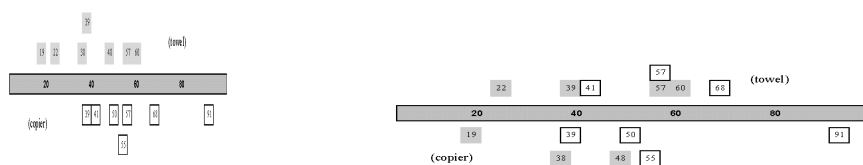
These principles, designed to build the bones that support our inferential meat, can be applied directly to learning about inference as well. In one activity (called *Orbital Express*), students test two rival designs for systems to deliver packages from Earth orbit. One design is a crumpled sheet of copier paper; another a similarly-crumpled paper towel. The test is to drop the “packages” repeatedly from orbit (standing on a chair) onto an earthbound target (a coin on the floor). Students measure how far each package is from the target when it comes to rest (that is, bounces count). The question is, which type of paper comes closer to the target?

This is a two-sample t situation. Instead, however, we use paper, a ruler, and computer-intensive methods to assess which design is better.

First, students represent their data concretely. They write each measurement (seven measurements for each design) on a slip of paper and place it next to a measuring strip taped to their table. Thus, they make a side-by-side plot of the distances at 100% scale. The scale is important: it helps students connect the experience to the data they’re working with.

Next, students construct their own measure. They have to devise a way to get a single number that tells them how much better (i.e., closer to the target) one design is than the other. Typically, students choose the difference in means or the difference in medians.

Suppose the difference in medians is 16 centimeters. Is 16 so large that it's unlikely to be due to chance? Instead of doing a t-test, we do a shuffle—first by hand, and then with the computer.

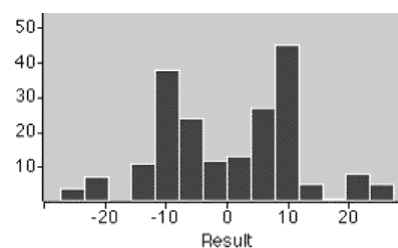


original data (difference of medians = 16 cm)

after shuffle (9 cm)

By hand, they take all 14 slips of paper, shuffle them, and deal them out to both sides. Then they place the cards next to the tape and calculate the measure again. By shuffling, we guarantee that there is no systematic difference between the sides of the tape. Nevertheless, the measure is not zero. Suppose we got a 9-cm difference in medians. We ask again, was our 16 cm by chance? We shuffle again, we get 12 cm. Students by now see that they need to do this many times. Software comes to the rescue, shuffling our data hundreds of times if necessary. The figure shows the distribution of 200 differences of medians.

Students see that if our test value lies well within the distribution of measures, we have no business claiming a difference. If, on the other hand, it would be an outlier when added to that distribution, we can legitimately claim that a measure that large is unlikely to arise by chance.



By using offline (paper) activities and software in concert, students are better able to come to grips with the slippery logic of the hypothesis test. By constructing their own measure and distribution, they will be better able to cope with *t* when they need to.

CONCLUSION

In this short paper, we have concentrated on one aspect of a statistical “skeleton”: firm knowledge of what distributions *are*, what they are *for*, and, when presented with a distribution, what it is a distribution *of*. Offline “paper” activities give students more

concrete experiences with distributions to build on and refer to. Flexible software lets students look at distributions of larger data sets in many different ways.

Finally, asking “what story do the data tell?” is really another way of asking that we help students look for meaning in the data at every step of the way. There is no need to present sterile and disembodied distributions; that path invites confusion and disinterest. Instead, by making things concrete and personal, by inviting conjecture, and by wrapping tasks in interesting contexts, we build not only interest but also a powerful suite of referents students can use as they build statistical muscle.