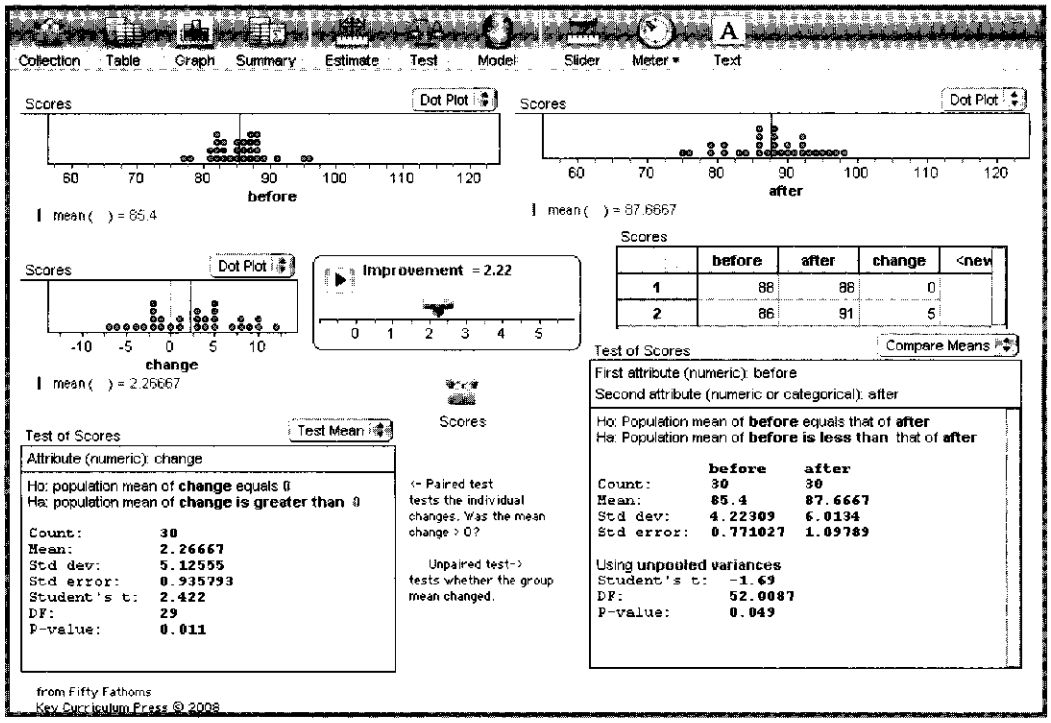
Demo 41: Paired Versus Unpaired

How a paired test gives a significant result more easily than its unpaired counterpart

Suppose you've given a pre-test and a post-test to a group of students and you want to see if your instruction had any effect. In your statistical toolbox for comparing two groups, you have both *paired* and *unpaired* techniques. In a paired test, you take the two scores for each individual student and find the difference; that is, you see how much each student's score changed. In the unpaired test, you test to see whether the "after" mean is significantly higher than the "before."

Which test should you use? In the situation in this demo, paired is best, because it's more powerful: You get a smaller *P*-value with the same data. Let's see how it looks.



What To Do

▷ Open **Paired Versus Unpaired.ftm.** It will look something like the illustration.

This is a complex-looking document, but there's really not a lot to it. You see the top of the data in the case table, middle right. In it, you can see three attributes: **before, after**, and **change**, representing the pre- and post-scores and the individual change in score. There's also a slider, **Improvement**, that controls the average amount of that change.

Then, in the lower left, is the "paired" test: We test whether the mean of **change** is greater than zero. Lower right, below the case table, is the "unpaired" test: We compare means to see if the mean of **after** is greater than that of **before.**

▷ Click on a point in any of the three dot plots. The same case will highlight in the other two graphs. Click on other points, too; study them to make sure you understand how the three attributes are related.

▷ Observe the test results. In the illustration, for example, the *P*-value for the paired (left-hand) test is smaller—more significant—than that of the unpaired (right-hand) test.

- ▹ Choose **Rerandomize** from the **Collection** menu. The points and the test results change. Again observe the difference in the *P*-values.
  - ⇨ The shortcut for **Rerandomize** is **⌘+Y** on the Mac or **Control+Y** in Windows.
- ▹ Move the **Improvement** slider, observing the effect on the graphs and on the test results.
- ▹ Find a value for the **Improvement** slider so that the *P*-value of the paired test is generally below 5%, while that of the unpaired test is generally above. (**Improvement = 1.5** is not bad.) **Rerandomize** repeatedly to get a sense of the variation.

At this point, you've been focusing on the numbers. Now step back and look at the graphs. Can you see the difference between the two sides? If so, you would see this: It's fairly clear that the left-hand **change** graph does not come from a population with a mean of zero (that is, the individual scores have generally improved), but it's not clear that the two right-hand graphs—**before** and **after**—come from populations with different means.

You can also think of it this way: If your **before** data are all spread out, but *every* student makes a small improvement, the mean of **after** will be higher, but not by much compared with the spread in the distribution of scores. So the unpaired test will have a small *t*-value. On the other hand, **change** will be tightly clustered around a number higher than zero—a number larger than its spread. So the paired test gets a large value for Student's *t* and a lower *P*-value. The paired test will be significant; whether it's meaningful is a different issue!

- ▹ Finally, once again select single points on the graphs to reinforce your understanding of their connection.

Questions

1. The **Improvement** slider should be roughly equal to the mean improvement of the scores. Is it?
2. Where can you find the mean improvement in the scores in the paired test?
3. Where can you find the mean improvement in the scores in the unpaired test? **Sol**

Challenges

4. Explain, as clearly as possible, how it can be that a paired test would give you better results than an unpaired test.
5. For the unpaired test, we compared two means. That makes sense. But for the paired test, we tested only a single mean. Explain why.
6. The formula for **before** is **randomBeta(4, 4, 70,100)**. Change the formula for **after** to be **randomBeta(4, 4, 70 + Improvement, 100 + Improvement)**. Now see what happens and explain it.
7. Suppose you're designing a study to see whether students' attitudes toward drug use change after they have seen a series of films. Your main measurement instrument is a questionnaire. What are the main arguments for and against having students put their names on it? **Sol**