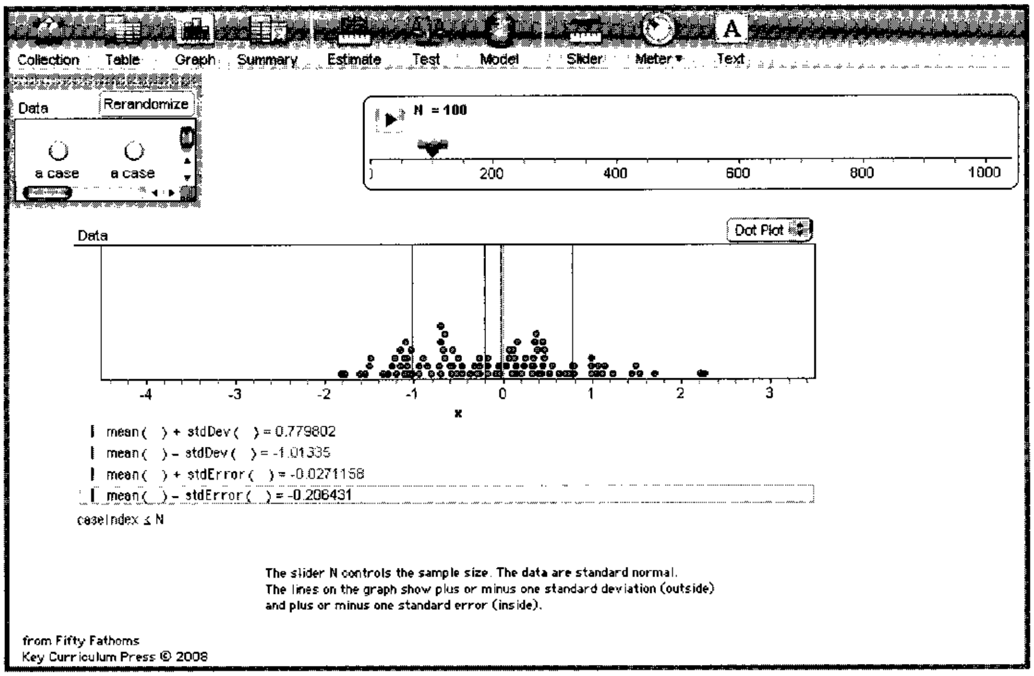
Demo 16: Standard Error and Standard Deviation

Getting a feel for the difference between standard deviation and standard error

It's easy to confuse standard deviation and standard error. Not only do they both start with *standard,* they're closely related. This informal demo should help you develop some intuition about the difference; Demo 17, "What Is Standard Error, Really?" goes into greater depth about standard error and helps lead toward understanding the *t*-statistic.



What To Do

▹ Open **SD and SE.ftm.** It should look like the illustration.

This relatively uncomplicated window contains a single collection, **data,** in the upper left. It has 1000 cases and one attribute, **x.** Fathom pulls these **x**-values randomly from a "standard" normal distribution—one with a mean of zero and a standard deviation of one. The graph shows the distribution of the first **N** cases, where **N** is controlled by the slider.

The graph also shows the positions and numerical values of the mean ± one standard *deviation* (the outside pair of lines) and the mean ± one standard *error* (the inside pair).

▹ Drag the slider **N** to the right. Note how the outside lines stay roughly in the same places—near plus or minus one.

▹ Do the same thing, paying attention to the positions of the inside lines—the ones that show how big the standard error is. Be sure to take **N** far below 100.

You should have seen that, though it may jump around a little, the standard deviation (SD) interval remains roughly constant, but the standard error (SE) interval gets larger as **N** decreases.

What is the difference between these two? Let's see:

- ▹ Reset **N** to 100 (you can edit the number in the slider and press **Enter, Tab,** or **Return**). Press the **Rerandomize** button in the collection (upper left). The points will change in the graph, and the lines will move.
- ▹ Repeat, noting what proportion of the time the SE interval encloses the true mean (0.00).
- ▹ Do the same thing, but with **N** = 400.

You should see that in both cases, the SE interval captures the true mean about two-thirds of the time. The SD interval, on the other hand, captures roughly two-thirds of the *data,* but almost always covers the true mean.

You can think of it this way: if you pick a point at random from the distribution, it is likely to be about 1 SD[1] from the mean. But if you pick a sample from the distribution and compute *its* mean, how far is *that* likely to be from the true mean? About 1 SE.

Extension

Let's study the SE interval more closely.

- ▹ Reset **N** to 100. Press the **Rerandomize** button.
- ▹ Repeat, noting the *width* of the SE interval.
- ▹ Record the full width of that interval at **N** = 25, 100, 400, and 900. Get several values at each **N.** (You could even enter these in a new Fathom collection with two attributes—**N** and **width**—and type in the values you recorded. Or you could write a formula and collect measures.)
- ▹ Note how, even though the interval jumps around, its width decreases (and becomes more consistent) as **N** increases. In particular, note that the width is roughly $2/\sqrt{N}$ —that is, 0.4, 0.2, 0.1, and 0.067.

The root-*N* dependence of standard error is the same as that of the spread in random walks (as in Demo 12, "How Random Walks Go as Root *N*"). This is no coincidence. In both cases, we're talking about how the spread of a sampling distribution decreases with the size of the sample. The random walk is about proportion; this one is about a "regular" mean.

[1]Depending on what you mean by a likely distance; we're speaking informally here. If you really mean root-mean-square distance, it's 1.00 SD by definition. On the other hand, the mean absolute distance for normal data is actually about 0.8 SD.