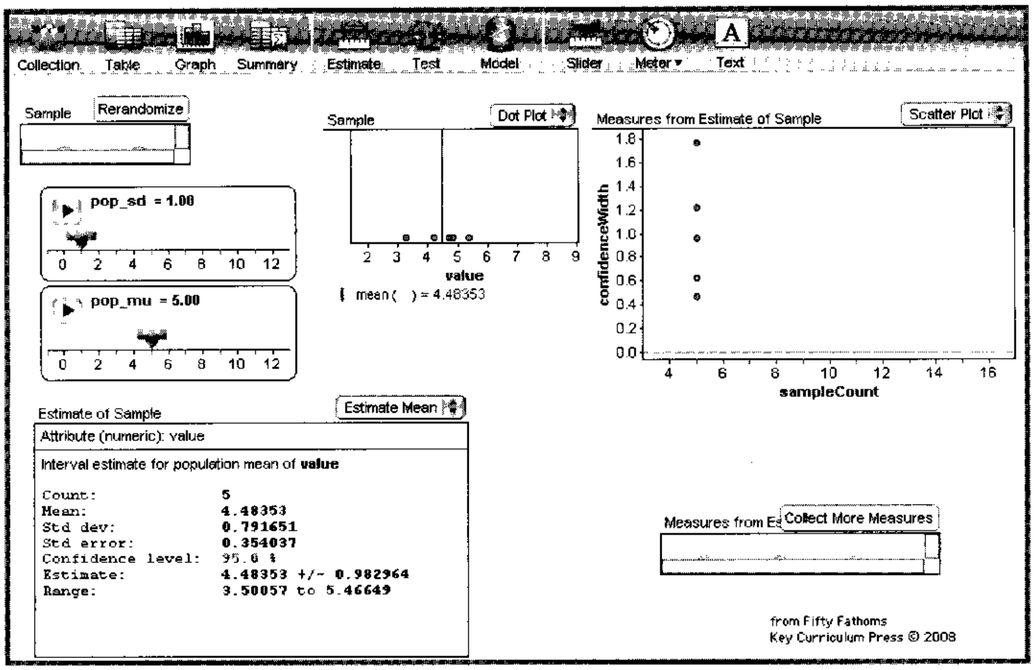
Demo 32: How the Width of the CI Depends on *N*

How the width of a confidence interval is inversely proportional to the square root of the sample size

In this demo, instead of making the distributions of means as we did in Demo 24, "How the Width of the Sampling Distribution Depends on *N*," we will start with a Fathom estimate, that is, we'll compute a confidence interval based on our sample. Then we'll track the width of the interval as the sample increases in size.



What To Do

Open **Width of CI depends on N.ftm**. It should look something like the illustration.

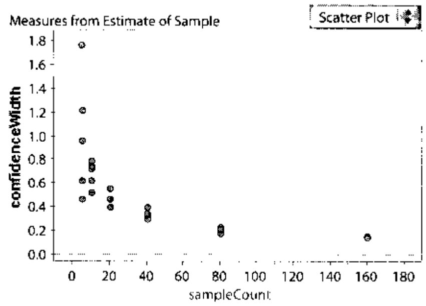
The collection **Sample** currently has five cases, drawn from a normal population with mean **pop_mu** and standard deviation **pop_sd.** They're graphed in the dot plot, top center. At the lower left, we have an estimate—a Fathom object that, in this case, computes a 95% confidence interval for the mean.

On the right-hand side, the measures collection and the graph above it collect and display information about repeated tests on new samples. So you can see that for a **sampleCount** of 5, there have been five tests. The width of their confidence intervals has ranged from about 0.4 to about 1.8. Note: **confidenceWidth** is the *half*-width of the entire interval.

- Press **Rerandomize** in the **Sample** collection a few times to see how the sample changes. The test changes too; see how the width of the confidence interval jumps around.
- Let's increase our sample size from 5 to 10. With the **Sample** collection selected, choose **New Cases** from the **Collection** menu. Add 5 cases for a total of 10. You'll see the middle dot plot update.
- Now let's record some confidence intervals. Press the **Collect More Measures** button in the measures collection at lower right. Fathom

makes five CIs from samples of 10 and adds their **confidenceWidth**s to the graph, upper right.

▷ Alternate between adding cases and collecting measures. Add 10 cases for a total of 20, 20 for a total of 40, 40 for a total of 80, and 80 for a total of 160. You don't have to do exactly these numbers. Be sure to **Collect More Measures** every time you add cases to the **Sample** collection. When you're done, your graph will look something like the one in the illustration.



Note: At this point in the demo, depending on experience, it's great if you can try different functions to fit these points. An inverse dependence and a decreasing exponential are popular candidates. For brevity, we omit the specific instructions. We can also use logarithms; see the extension.

▷ Let's make a function and see if the one-over-root-*N* dependence is at work here. Select the graph by clicking on it once. Then choose **Plot Function** from the **Graph** menu. The formula editor appears.

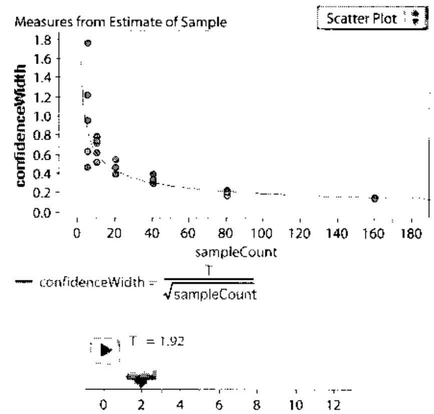
▷ Enter **1/$\sqrt{\textbf{sampleCount}}$** and press **OK.** The graph will miss, but it's about the right shape.

▷ Let's make a variable parameter so that we can fix the function. Drag a slider from the shelf and put it in the blank area below the graph. Edit its name **V1** to be **T.**

▷ Double-click the formula for **confidenceWidth** at the bottom of the graph to open the formula editor again.

▷ Put **T** in the numerator. Again, press **OK** to close the editor.

▷ Use the slider to change **T** so that the curve fits the data. You may want to zoom in to the cluster of points over 160 to make sure you hit them. Your graph and slider will look something like the ones in the illustration.



It is good to stop here and reflect on what an inverse root function means: since 160 is four times 40, the width of the confidence interval will be $1/\sqrt{4}$, or one-half the size. Similarly, by increasing the sample size from 5 to 80, you reduce the confidence width by a factor of 4.

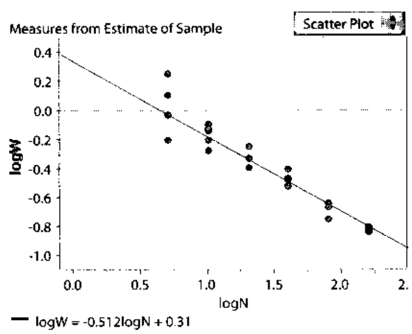
Questions

1. How large a sample would we need to reduce the "margin of error" to 0.01? That is, how many points do we need in order to get **confidenceWidth** as low as 0.01? **Sol**
2. What would be different if we changed **pop_mu**?
3. What would be different if we changed **pop_sd**?
4. What would be different of we changed the confidence level in the test from 95% to, say, 80%?
5. What would be different if we did 10 tests at every sample size instead of just 5?
6. What does it mean that the values of **confidenceWidth** get closer together as **sampleCount** increases? **Sol**

Extension

If you didn't know about root-N, you could use logarithms to discover it.

- Double-click the measures collection to open its inspector. Click the **Cases** tab to go to the **Cases** panel.
- Make two new attributes, **logN** and **logW.**
- Give them formulas[3] so that they are the logarithm of **sampleCount** and the logarithm of **confidenceWidth,** respectively.
- Then plot **logW** against **logN** and fit a movable line.[4] (Ignore the points with a sample size of 5; see the challenges, below.) You'll get a graph like the one in the illustration. Note the slope; figure out what it means.



Challenges

7 Questions 1–6 each imply an activity; try them. For example, redo this demo with **pop_sd** equal to 2 instead of 1. What happens?

8 It may not have been clear in the original, curvy graph, but it's obvious in the log-log graph above that the tests with **sampleCount = 5** give a larger **confidenceWidth** than you would expect based on the pattern from the other sample sizes. Explain why.

Tautology Alert

The experienced reader may cry "foul!" after reading and exploring this demo. We discovered that if you make confidence intervals for different sample sizes (all drawn from the same population), the width of the CI is pretty closely inversely proportional to the square root of the sample size. In fact, the fit is so good, you may be suspicious.

Your suspicions will be confirmed when you realize that Fathom computes the width of the CI to be

$$w = t^* \frac{s}{\sqrt{N}}$$

where t^* is the relevant t-value and s is the sample standard deviation. So of course it will have a one-over-root-N dependence. So, why include this demo? Several reasons:

- Most importantly, we wanted to approach the issue using a different technique than we did in Demo 24, "How the Width of the Sampling Distribution Depends on N." Sure, it's the same principle, but we do not always recognize that the same principle applies when we dress it up in different clothes.
- Secondly, knowing that there's a root N in the denominator of the function doesn't mean we have a feel for what that really means. So actually making samples—with real variation—and testing them repeatedly makes good sense.
- Finally, and most subtly, since t^* depends on n, the width is *not* in fact inversely proportional to the square root of the sample size; the departure from that rule is exactly why we need t instead of just z.

[3]For example, **log(sampleCount).**

[4]To get a movable line, first select the graph, then choose **Movable Line** from the **Graph** menu. Drag parts of the line to change its slope and intercept; the equation appears at the bottom of the graph.