

Are You Confident?

by – Dennis Ivany

Activity overview

Problem 1 - *A Brief Review of the Normal Distribution. This is intended as a review only, and students should have had significant prior exposure to the concepts presented.*

Problem 2 *develops the meaning of a confidence interval. Included are some opportunities for students to visually explore confidence intervals. A sampling distribution of the sample mean for a known population is created and students can drag line segments representing 95%, 90%, and 99% intervals to center them over various sample means for the population, thus getting a feel for when an interval captures the population mean.*

Concepts

Normal Distribution

Random Sampling

Sampling Distribution of the Sample Mean

Confidence Intervals

Teacher preparation

*It might be helpful for teachers to replicate the activity by placing a drawing of the sampling distribution of the sample mean for problem 2 on the wall. Sample means can be generated on the TI-Nspire using the **mean(randNorm(** command. Dots can be placed on sticky notes and placed on the sampling distribution of the sample mean, then string of appropriate length for each confidence interval can be centered over the sample means to see whether they capture the population mean.*

Classroom management tips

It is best if students are familiar with the normal distribution prior to beginning problem 2. This would usually be accomplished in a prior course. The material in problem 1 is presented only as a refresher.

TI-Nspire Applications


Graphs and Geometry

Notes

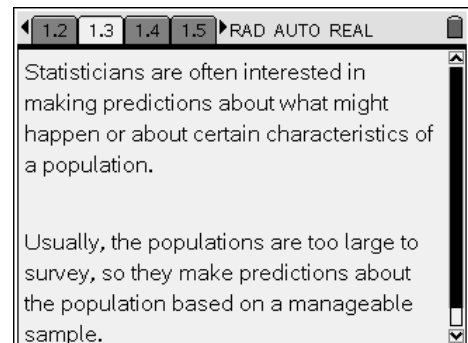
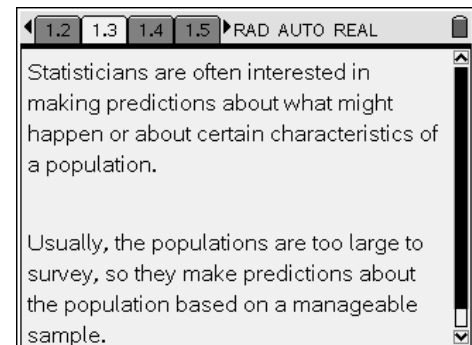
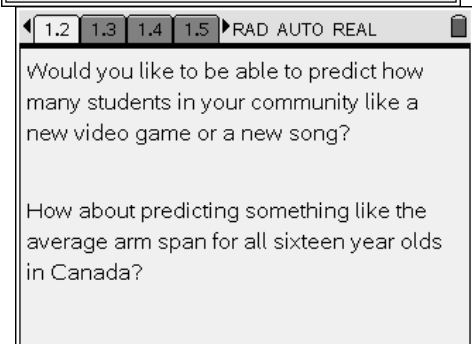
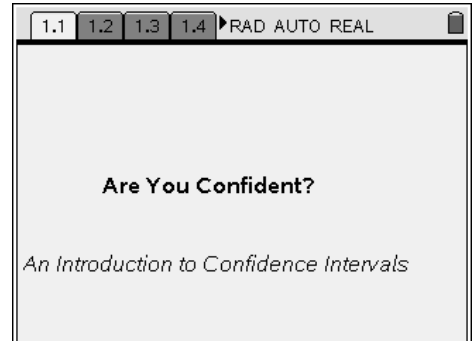
Lists and Spreadsheets

Step-by-step directions

Open the document stats_Confint_Ivany.

Use  to move between slides.

Read the first four slides.



Are you confident?

by: Dennis Ivany

Grade level: secondary

Subject: Statistics

Time required: 45 to 90 minutes

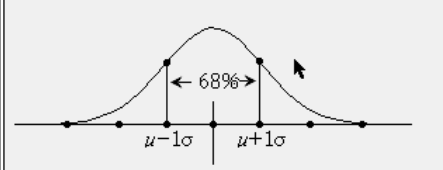
Materials: TI-Nspire

The remaining slides up to 1.18 are intended as a refresher of the normal distribution. If you do not need them, go to slide 2.1 under Problem 2.

1.2 1.3 1.4 1.5 RAD AUTO REAL

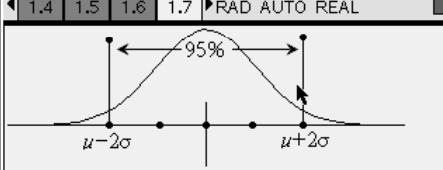
Review of the Normal Distribution
(optional)

1.3 1.4 1.5 1.6 RAD AUTO REAL



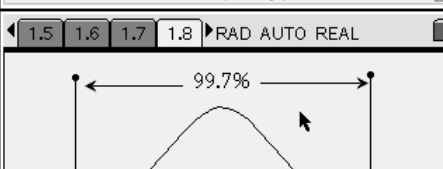
In a normal distribution, about 68% of the data lies within 1 standard deviation of the population mean.

1.4 1.5 1.6 1.7 RAD AUTO REAL



About 95% of the data lies within about 2 standard deviations of the population mean (actually, it's closer to 1.96 than 2, but we'll use 2 for simplicity).

1.5 1.6 1.7 1.8 RAD AUTO REAL



99.7% of the data lies within about 3 standard deviations of the population mean

Are you confident?

by: Dennis Ivany

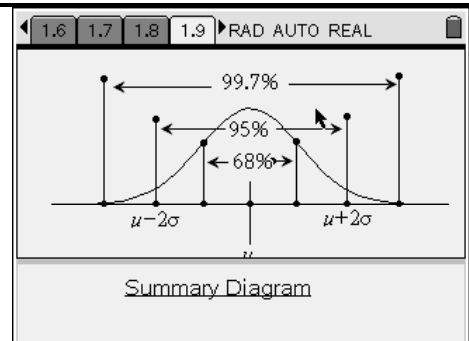
Grade level: secondary

Subject: Statistics

Time required: 45 to 90 minutes

Materials: TI-Nspire

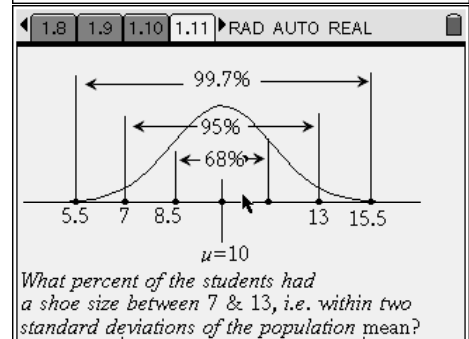
A summary of common intervals on a normal distribution.



Have students write down the mean and standard deviation before proceeding to the next slide.

Now, let's apply this to a problem involving shoe sizes among the Canadian men. Let's assume we know that the mean shoe size is 10, with a standard deviation of 1.5. We can label our normal distribution as follows:

The answer for the question will be found on the next slide.



Type your answer by the question.

Press **tab** to move to the **Answer** line and press **enter** to see the answer.

Question

What percent of the students had a shoe size between 7 and 13, i.e. within two standard deviations of the population mean?

Answer

Are you confident?

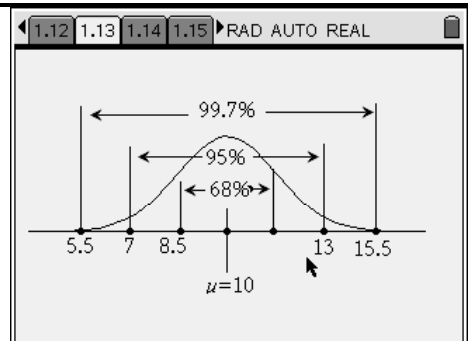
by: Dennis Ivany
 Grade level: secondary
 Subject: Statistics
 Time required: 45 to 90 minutes

Materials: TI-Nspire

A repeat of the earlier diagram to help you with the questions on the next slide.

Press **tab** to move to the **Answer** line and press **enter** to see the answer.

Students should take note by copying the information on slides 1.16 and 1.17 into their notes.




1.12 1.13 1.14 1.15 RAD AUTO REAL

Question

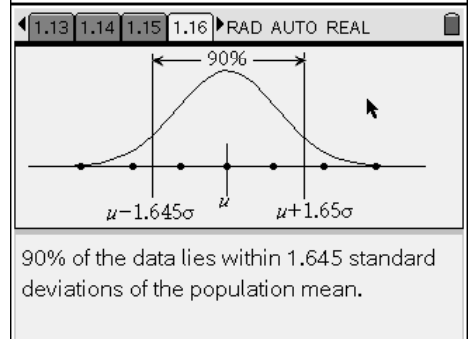
What percent of the students had shoe sizes between:

A) 8.5 and 11.5?
 B) 8.5 and 15.5?

Answer 

1.12 1.13 1.14 1.15 RAD AUTO REAL

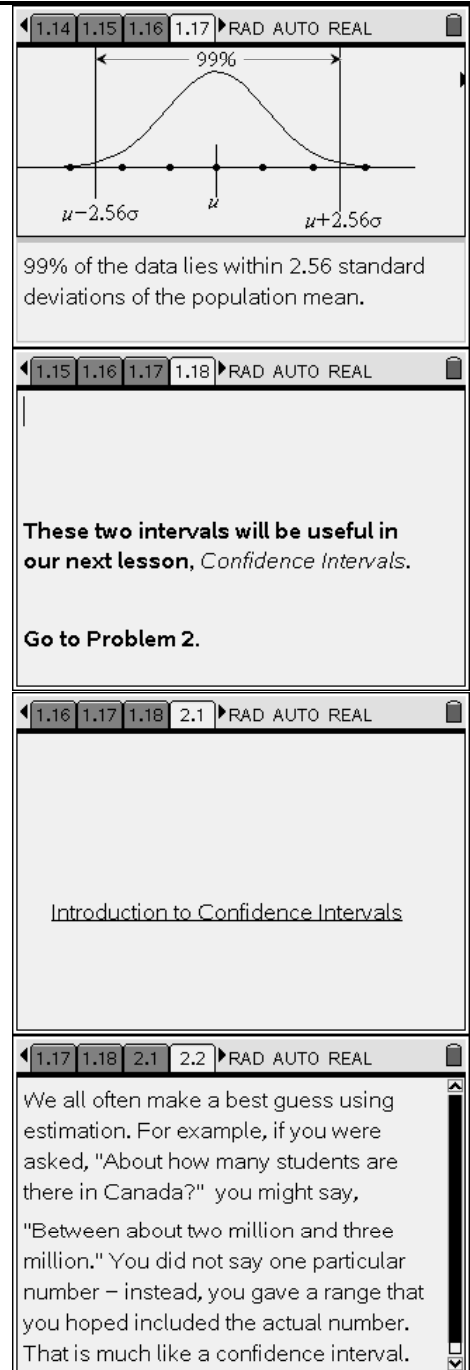
Some other useful intervals:



Are you confident?

by: Dennis Ivany
 Grade level: secondary
 Subject: Statistics
 Time required: 45 to 90 minutes

Materials: TI-Nspire



1.14 1.15 1.16 1.17 RAD AUTO REAL

99%

$\mu - 2.56\sigma$ μ $\mu + 2.56\sigma$

99% of the data lies within 2.56 standard deviations of the population mean.

1.15 1.16 1.17 1.18 RAD AUTO REAL

These two intervals will be useful in our next lesson, *Confidence Intervals*.

Go to Problem 2.

1.16 1.17 1.18 2.1 RAD AUTO REAL

Introduction to Confidence Intervals

1.17 1.18 2.1 2.2 RAD AUTO REAL

We all often make a best guess using estimation. For example, if you were asked, "About how many students are there in Canada?" you might say, "Between about two million and three million." You did not say one particular number – instead, you gave a range that you hoped included the actual number. That is much like a confidence interval.

This is where the introduction to confidence intervals really begins.

Students read the notes on slides 2.2 through 2.5.

Are you confident?

by: Dennis Ivany
 Grade level: secondary
 Subject: Statistics
 Time required: 45 to 90 minutes

Materials: TI-Nspire

◀ 2.1 2.2 2.3 2.4 ▶ RAD AUTO REAL

The problem with your guess is that it was very informal and it is hard to know whether you were even close.

Statisticians use a formal process that involves taking a sample from the population and then making a guess based on a formula. We'll try to see a picture of the process instead.

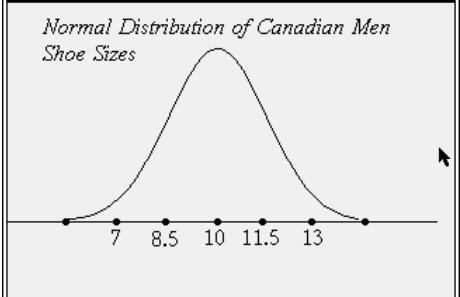
◀ 2.1 2.2 2.3 2.4 ▶ RAD AUTO REAL

As an example, let's pretend that the government knows the average shoe size of Canadian men is size 10, with a standard deviation of 1.5.

We will assume these shoe sizes are normally distributed.

◀ 2.2 2.3 2.4 2.5 ▶ RAD AUTO REAL

Normal Distribution of Canadian Men Shoe Sizes



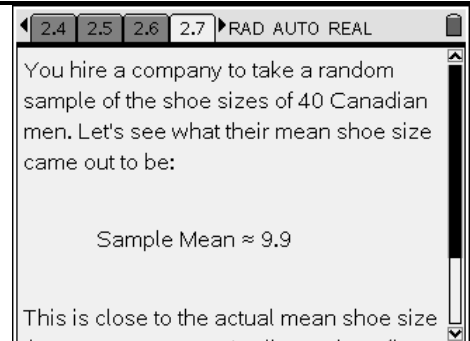
◀ 2.3 2.4 2.5 2.6 ▶ RAD AUTO REAL

Now let's also pretend that the government will not tell anyone what they know, but you need to know because you own a shoe store and you want to decide on what range of sizes to order.

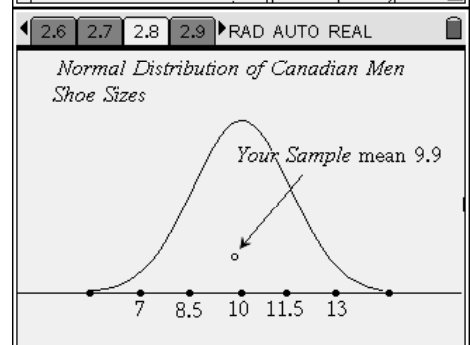
Remember, you want to make money so you want to buy the most popular range of sizes.

Putting the problem in context.

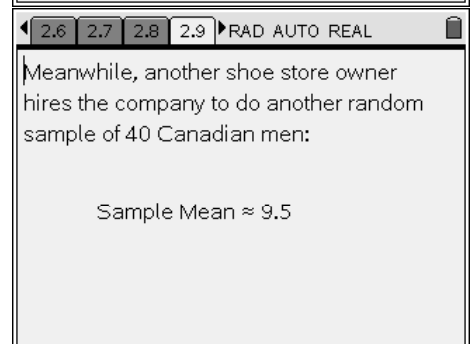
Write down this sample mean.



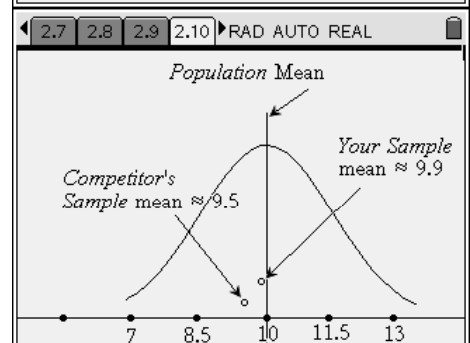
Your sample mean plotted in relation to the actual population.



Your competitor's sample mean.



In this case, both sample means are pretty close to the population mean.



Are you confident?

by: Dennis Ivany
 Grade level: secondary
 Subject: Statistics
 Time required: 45 to 90 minutes

Materials: TI-Nspire

Scrolling down through this spreadsheet will reveal the sample means from the 100 stores.

2.8 2.9 2.10 2.11 RAD AUTO REAL

For some strange reason, shoe stores keep hiring the company to sample 40 Canadian men shoe sizes. In total, 100 store owners (including you) do this.

The 100 Sample Means are shown in the column "sample.means" on the next page.

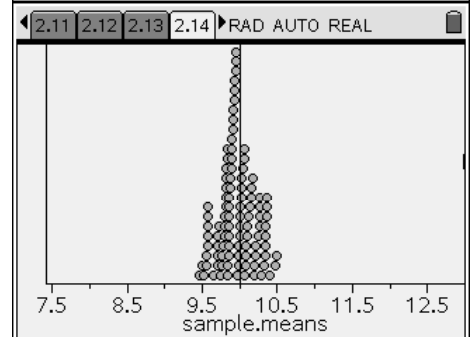
2.9 2.10 2.11 2.12 RAD AUTO REAL

A	sample...	B	C	D	E	F
1	9.46548					.0
2	9.49254					.0
3	9.52016					.0
4	9.55363					.0
5	9.5609					.0
A7	9.465475254					

Notice how this dot plot is shaped much like a normal distribution.

2.10 2.11 2.12 2.13 RAD AUTO REAL

The next page shows the sample means as a dot plot tightly clustered around the population mean shoe size of 10.



Are you confident?

by: Dennis Ivany
 Grade level: secondary
 Subject: Statistics
 Time required: 45 to 90 minutes

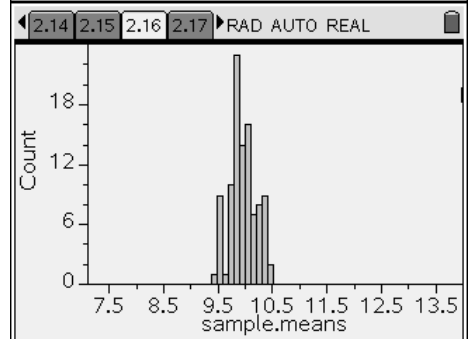
Materials: TI-Nspire

Press **tab** to move to the **Answer** line and press **enter** to see the answer.

◀ 2.14 2.15 2.16 2.17 ▶ RAD AUTO REAL

The next page shows a histogram of the sample means.

Notice that is centered on the population mean shoe size of 10.



◀ 2.14 2.15 2.16 2.17 ▶ RAD AUTO REAL

Question

Which best describes the shape of the histogram?
 A) Uniform B) Bell

Answer ⬆

◀ 2.15 2.16 2.17 2.18 ▶ RAD AUTO REAL

Statisticians have determined that if we were able to repeat the sampling process an infinite number of times, the distribution of sample means would itself be normally distributed!

So what?

Are you confident?

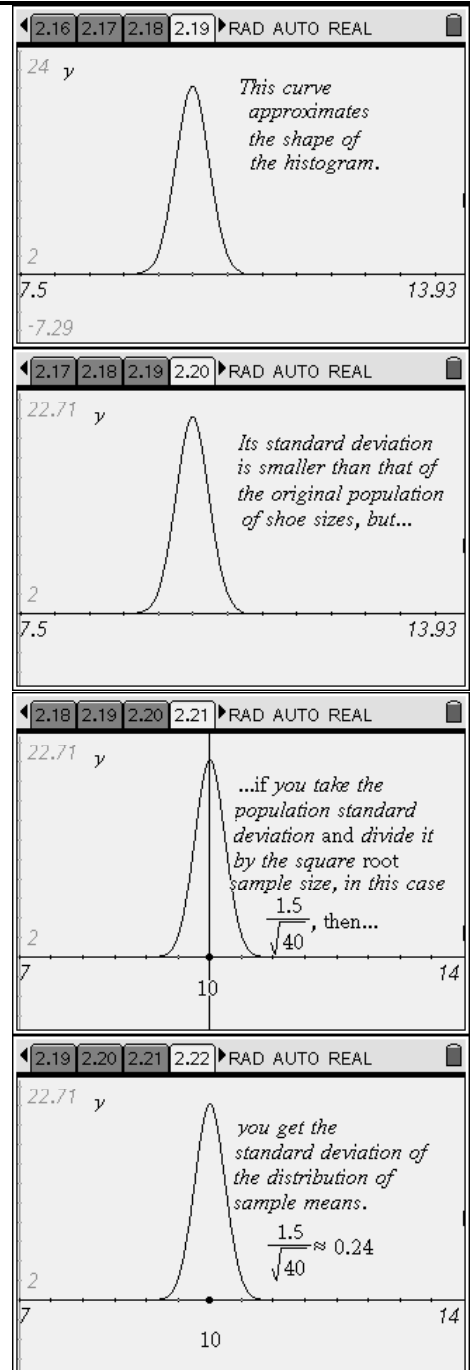
by: Dennis Ivany

Grade level: secondary

Subject: Statistics

Time required: 45 to 90 minutes

Materials: TI-Nspire



Are you confident?

by: Dennis Ivany

Grade level: secondary

Subject: Statistics

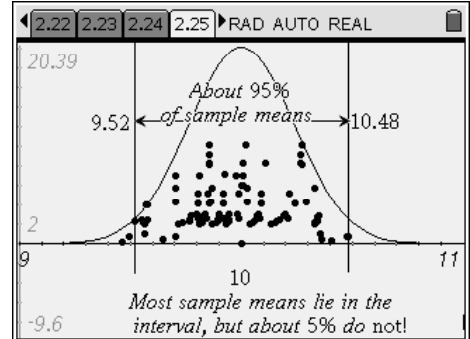
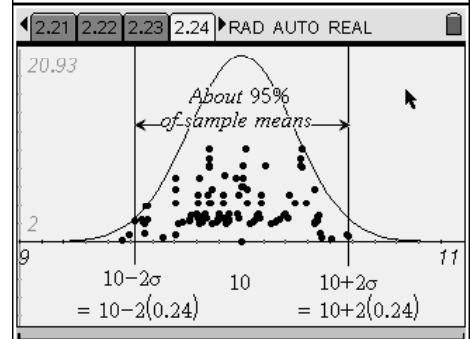
Time required: 45 to 90 minutes

Materials: TI-Nspire

Note: The height of a point above the x-axis has no relevance and has been randomly assigned.



2.20 2.21 2.22 2.23 RAD AUTO REAL

The scale on the horizontal axis has been changed to magnify the view of the distribution of sample means.





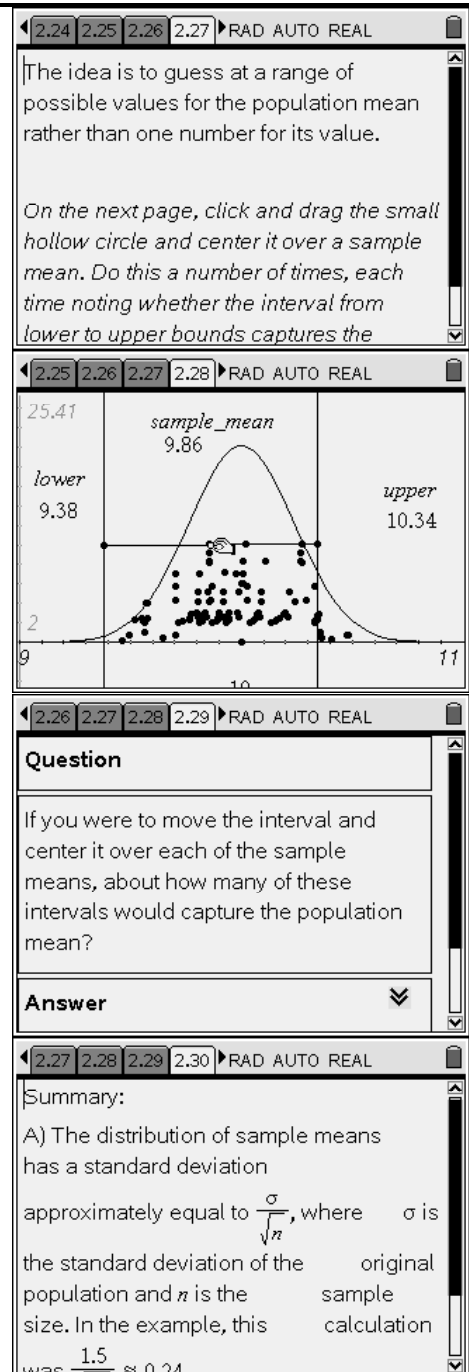
2.23 2.24 2.25 2.26 RAD AUTO REAL

In the real world, though, we can only take one sample. How do we make predictions based on one sample?

Place the hand on the line segment or on the open circle then press   and drag the line segment around.

Center it over a sample mean. Is the population mean of 10 captured within the interval between the lower and upper bounds of the line segment? Try to find all the sample means where the interval does NOT capture the population mean.

Press  to move to the **Answer** line and press  to see the answer.



The image shows three screenshots from a TI-Nspire calculator interface. The top screenshot shows a normal distribution curve with a mean of 9.86 and a standard deviation of 2. The lower bound is 9.38 and the upper bound is 10.34. The x-axis is labeled with 9, 10, and 11. The y-axis is labeled with 2 and 25.41. The middle screenshot shows a question: "If you were to move the interval and center it over each of the sample means, about how many of these intervals would capture the population mean?" The bottom screenshot shows a summary of the distribution: "A) The distribution of sample means has a standard deviation approximately equal to $\frac{\sigma}{\sqrt{n}}$, where σ is the standard deviation of the original population and n is the sample size. In the example, this calculation was $\frac{1.5}{\sqrt{25}} \approx 0.3$." The calculator interface includes a menu bar with options like 2.24, 2.25, 2.26, 2.27, RAD, AUTO, REAL.

Are you confident?

by: Dennis Ivany
 Grade level: secondary
 Subject: Statistics
 Time required: 45 to 90 minutes

Materials: TI-Nspire

◀ 2.28 2.29 2.30 2.31 ▶ RAD AUTO REAL

B) About 95 of every 100 (or 95% of them)sample means was within 2 of these standard deviations from the population mean.

C) If we construct a segment that stretches from 2 standard deviations below the mean to 2 standard deviations above, we

◀ 2.29 2.30 2.31 2.32 ▶ RAD AUTO REAL

D) We can expect that about 95 of every 100 intervals will capture the population mean.

◀ 2.30 2.31 2.32 2.33 ▶ RAD AUTO REAL

Experimenting with other Intervals

◀ 2.31 2.32 2.33 2.34 ▶ RAD AUTO REAL

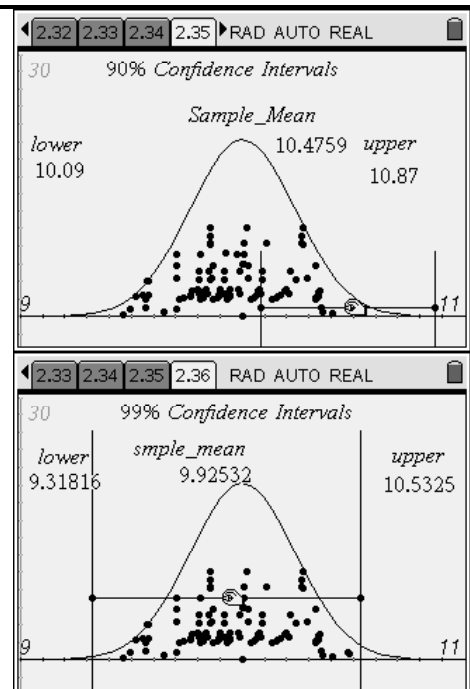
A 90% interval is not as long as the 95% interval for the same sample size from the same population, and so we would have less confidence in how often it captures the population mean.

Experiment with the 90% interval for our shoe size example on the next page.

Notice how the 90% interval is shorter. We can expect that fewer intervals will capture the population mean than when we used 95% intervals.

Again, drag the line segment around. Take note of the locations where the interval does not capture the population mean.

The 99% interval is longer. It should capture the population mean when centered on 99 of every 100 sample means.



Assessment and evaluation

- *Journal entry explaining what a confidence interval is.*
- *Use an "exit slip" – ask students to identify one or two key points they really understood or are still confused about. Each student must write about her/his key points on a slip of paper, or a form provided by the teacher, and pass it in before permission is granted to leave the classroom for the day.*
- *Have each student find ten of the 95% confidence intervals in the activity (each student should determine her/his own particular ten and not deliberately try to use the same 10 as someone else), and write down the number of intervals that capture the population mean. Have the class find the average number intervals that capture the population mean. What percent of the time does an interval capture the population mean? (The answer should be close to 95%).*