THE CONFIDENCE INTERVAL

6:27 / 6:29 1.0x

SPEAKER: MICHAEL J. MAHOMETA, Ph.D.

Confidence.

What does it mean in the context of statistics?

It all has to do with a sample, and its relationship to the population

it was drawn from.

We know that there is a thing called Sampling Error

- that the mean of each sample we draw from a population will not necessarily have the population mean.

But will it be close?

And can we define a range of values - based on our sample data

- that point to where the population mean should be?

Help

The thing that my students tell me that they have a

really hard time understanding is Confidence Intervals.

So I'd like to show you a simulation that will hopefully

demonstrate just what we mean by confidence when

we use it in the contact of statistics.

We know that for any population, we can draw a sample of any size,

and that, naturally, there will be what's called Sampling Error.

That if we draw sample after sample after sample from the same population,

the mean of those samples will NOT actually be equal.

They'll each be a little different.

Now, not only will the means be different,

but the Standard Deviations for each sample will be different as well.

Help

I've got some simulated population data, with a specific population mean,

but I'm not going to show it to you.

Instead we'll use it to draw samples from.

Let's use it to draw 20 samples, each with a sample size of 15.

Now let's animate a plot for each of the 20 samples.

Here's the first sample, with a mean of 18.57.

Now, I've also drawn a 95% Confidence Interval around the mean.

Here's the next sample with a mean of 18.47 and its Confidence Interval.

And now the third sample with a mean of 21.42 and its Confidence Interval.

Let's speed this up a bit.

We see that it's true - each sample has a slightly different mean

and Standard Deviation (as shown through a confidence interval).

And something very interesting has happened.

Most of the Confidence Intervals overlap the central area of the plot.

Can we see where the majority of them overlap?

Maybe the use of a horizontal line will help.

Here's the horizontal line at 18, and we miss two means.

Here at 19 we miss one, and, at 20, we miss one again.

But at now 21, we miss two.

Help

So, knowing that each sample mean will bounce around the actual population

mean, if we had to guess, where do you think the actual population

mean might be for these samples?

I'd guess somewhere around 19 or 20.

Well here's the actual population mean; it's 20.

That means only one sample's 95% Confidence

Interval missed the actual population mean.

It's the sample right here.

And here's the cool thing - a little bit of quick math

and we see that 19 out of 20 is - you guessed it - 95%.

Do you think that it's a coincidence that these two numbers match?

95% of samples catch the population mean,

and we drew a 95% Confidence Interval around each mean?

No, it's no coincidence at all.

Let's take a look at the same samples, but

Help

now,

instead of calculating 95% Confidence Intervals,

we'll calculate 85% Confidence Intervals.

The population mean is still 20.

And what do you notice about the samples and their Confidence Intervals?

The Confidence Interval range is smaller - closer to the mean of the sample.

AND, now there are three samples that don't catch the population mean.

Again, it's no coincidence that 17 out of 20 is 85%.

And that's the key to understanding Confidence Intervals:

if we draw an infinite number of samples of the same size,

and we calculate each sample's mean and particular Confidence Interval, then

that Confidence Interval value of those samples

will actually "catch" the true population mean value. if we calculate

a 95% Confidence Interval value, we'll actually have 95% of the samples

"catch" the true population mean.

If we reduce the Confidence Interval to 85%, then only 85% of those sample

Help

Confidence Intervals will actually "catch" the true population mean value.

This is why we say, when using a 95% Confidence Interval,

that we are 95% confident that the true population falls

within our range of these values.

We're 95% confident, because we know that 5% of our sample Confidence

Intervals will in fact miss the true population value.

In the end, it's all a matter of balance.

Balance between Confidence and Precision.

The more confident I am, the less precise

my range is the larger the spread around my sample mean.

The more precise I am, by my smaller Confidence Interval, the less confident

I am that my particular mean catches the actual population mean.

Help

Comprehension Check

1. The average resting pulse of adult women is about 74 beats per minute. The standard deviation in the population is believed to be 13 beats per minute. The distribution of resting pulse rates is known to be skewed right.

(1/1 point)

1a. You take a sample of 36 women from this population and obtain a sample mean of 77 beats per minute. What is the shape of the sampling distribution from which your mean comes?

- (●) The sampling distribution will be nearly Normal because the sample size of 36 is sufficiently large. ✔
- ( ) The sampling distribution will be skewed right because the population is skewed right.
- ( ) The sampling distribution will not be Normal because the sample size is too small.

Check | Show Answer

(1/1 point)

1b. What value should be at the center of the sampling distribution?

- 70 beats per minute
- 74 beats per minute
- 77 beats per minute

Help

Check Hide Answer

(1/1 point)

1c. How much variation should we expect in sample means given a sample size of 36? In other words, what is the value of the standard error?

- 0.37
- 2.17
- 2.90
- 7.29

Check Hide Answer

(1/1 point)

1d. What is the z-score of our sample mean in the sampling distribution given sample sizes of 36? *(Round to 2 decimal places.)*

Help

- ( ) -1.19
- ( ) 0.23
- (•) 1.38 ✔
- ( ) 2.54

Check Hide Answer

(1/1 point)

1e. What is the probability of observing our sample mean, or one that is larger, when the true population mean is 74 beats per minute?

- ( ) 0.0002
- (•) 0.0838 ✔
- ( ) 0.2146
- ( ) 0.9127

Check Hide Answer

1f. Construct a 95% confidence interval around your sample mean of 77 beats per minute using this formula:

$$\text{Mean} = 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

Help

(1/1 point)

What are the bounds of the 95% confidence interval?

- (59.40, 92.40)
- (69.45, 78.75)
- (72.75, 81.25) ✔
- (74.15, 75.85)

Check | Hide Answer

(1/1 point)

1g. Does the 95% confidence interval give you results that are consistent or inconsistent with your earlier z-score?

- (•) Consistent. The 95% confidence interval shows that the population mean from which the sample was drawn could be 74. Similarly, the z-score of the sample mean places it clearly within the expected sample mean values for a known population mean of 74. ✔
- ( ) Inconsistent. The 95% confidence interval includes 74 as a possible sample mean, but the z-score of our sample mean is extreme. It is unlikely that our sample mean would be seen if the true population average is 74 beats per minute, as claimed.
- ( ) There is no relationship between the sampling distribution and the confidence interval, so we cannot answer this question.

Check | Hide Answer

Help

---

2. Consumer reports tested 14 brands of vanilla yogurt and found the following numbers of calories per serving:

180	200	190	230	80	160	170
130	140	220	110	120	100	170

(1/1 point)

2a. What is the sample mean? *(Round to 2 decimal places.)*

157.14

Help

157.14

Answer: 157.14

Check Hide Answer

(1/1 point)

2b. If the standard deviation reported by the yogurt industry is 48.5 calories, how much variation should we expect between sample means for samples of size 14? *(Round to 2 decimal places.)*

12.96

12.96

Answer: 12.96

Check Hide Answer

(1/1 point)

2c. What is the margin of error if the standard deviation is 48.5 and the samples are of size 14? *(Assume 95% confidence and round to 1 decimal place.)*

25.4

25.4

Answer: 25.4

Check Hide Answer

Help

(2/2 points)

2d. What is the 95% confidence interval for the average calorie content of vanilla yogurt? *(Round to 1 decimal place.)*

Lower Bound

131.7

131.7

Answer: 131.7

Upper Bound

182.5

182.5

Answer: 182.5

Check Hide Answer

(1/1 point)

2e. The FDA reports that the average caloric content of vanilla yogurt is 150 calories. Does your confidence interval support or refute this claim?

Help

- (•) Support ✔
- ( ) Does not support

Check Show Answer