Margin of Error

Whose error is it?
Why do we care?

Lesson Objectives

- Learn the meaning of “margin of error,” or “m.o.e.”
- Learn how to calculate the m.o.e. for two situations: when the true population std. dev., \( \sigma \), is known and when it is unknown.
- Learn how to use the m.o.e. to construct a confidence interval.

“Margin of Error” for estimating the True Mean of a population.

m.o.e. at 95% confidence =

“The amount that when added and subtracted to the true population mean will define a region that will include the middle 95% of all possible \( \bar{X} \)-bar values.”

Illustration of “Margin of Error”

\[ \mu = \text{true population mean} \]

Find the interval around the mean where 95% of all possible sample means will lie.

\[ m.o.e. = 1.96 \left( \frac{\sigma}{\sqrt{n}} \right) \]

Margin of Error for 95% confidence:

\[ m.o.e. = 1.96 \left( \frac{\sigma}{\sqrt{n}} \right) \]
General form for “margin of error” when \( \sigma \) is known:

\[
m.o.e. = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}
\]

where \( Z_{\alpha/2} \) is the appropriate percentile from the standard normal distribution, i.e., the Z table.

Explanation of symbol:
\( Z_{\alpha/2} \) cuts off the top tail at area = \( \alpha/2 \)

Estimation of symbol:
\( Z \sim N(0,1) \)

### Examples

<table>
<thead>
<tr>
<th>Amount of confidence</th>
<th>Area in each tail</th>
<th>Table value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1 - \alpha )</td>
<td>( \alpha/2 )</td>
<td>( Z_{\alpha/2} )</td>
</tr>
<tr>
<td>.98</td>
<td>.0100</td>
<td>2.33</td>
</tr>
<tr>
<td>.95</td>
<td>.0250</td>
<td>1.96</td>
</tr>
<tr>
<td>.90</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>.80</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

### (1-\( \alpha \))100% Confidence Interval:

Point estimate \( \pm \) m.o.e.

\( (1-\alpha)100\% \) is the amount of confidence desired.

\( \alpha = .05 \) risk

### Example 1

Estimate with 98% confidence the mean gallons of water used per shower for Dallas Cowboys after a game if the true standard deviation is known to be 10 gallons. The sample mean for 16 showers is 30.00 gal.

\[
m.o.e. = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}
\]

Example 1, continued . . .

98% confidence interval:

Point estimate \( \pm \) m.o.e.
Example 1, Statement in the L.O.P.

I am 95% confident that the true mean gallons of water used per shower for Dallas Cowboys after a game will fall within the interval 24.175 to 35.825 gallons.

A statement in L.O.P. must contain four parts:
1. amount of confidence.
2. the parameter being estimated in L.O.P.
3. the population to which we generalize in L.O.P.
4. the calculated interval.

What do we do if the true population standard deviation \( \sigma \) is unknown?

- Replace \( \sigma \) with \( \bar{s} \).
- replace \( z \) with \( t_{\alpha/2, n-1} \).

General form for “margin of error” when \( s \) is UN-known:

\[
m.o.e. = t_{\alpha/2, n-1} \sqrt{\frac{\bar{s}^2}{n}}
\]

where \( t_{\alpha/2, n-1} \) is the appropriate percentile from the \( t \)-distribution.

Explanation of symbol:

\( t_{\alpha/2, n-1} \) cuts off the top tail at area \( \alpha/2 \)

Use the \( t \)-table to find the value.

Comparison of “Z” and “t”

<table>
<thead>
<tr>
<th></th>
<th>( Z )</th>
<th>t-dist.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell shaped, symmetric</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean</td>
<td>( = 0 )</td>
<td>( = 0 )</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>( = 1 )</td>
<td>( &gt; 1 )</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>“( \infty )”</td>
<td>“( n-1 )”</td>
</tr>
</tbody>
</table>

As “\( n-1 \)” increases, “\( t_{n-1} \)” approaches “\( Z \)”.

Want 95% CI, \( n = 20 \),
\[
\alpha/2 = .025
\]
\[
d.f. = 19
\]
\[
t_{.025, 19} = 2.093
\]
Want 98% CI, 
\[ n = 33, \quad \alpha/2 = \frac{0.01}{2} = 0.005 \]
\[ d.f. = \frac{n}{2} = \frac{33}{2} = 16.5 \]
\[ t \approx 2.131 \]
\[ \text{Margin of Error} = t \times \frac{s}{\sqrt{n}} = 2.131 \times \frac{10.4}{\sqrt{33}} \approx 3.66 \]
Confidence vs. Probability

- **BEFORE** a sample is collected, there is a 95% probability that the future to be computed sample mean will fall within m.o.e. units of μ.
- **AFTER** the sample is collected, the computed sample mean either fell within m.o.e. units of μ or it did not. After the event, it does not make sense to talk about probability.

**Analogy:** Suppose you own 95 tickets in a 100-ticket lottery. The drawing was held one hour ago, but you don’t know the result. P(win) = 0 or 1, but you are very CONFIDENT that you have won the lottery.

Interpretations of the Confidence Interval for μ

- If we took many, many samples of size n, and calculated a confidence interval for each, then I would expect that 95% of all these many intervals would contain the true mean, and 5% would not.

What sample size is needed to estimate the mean mpg of Toyota Camrys with an m.o.e. of 0.2 mpg at 90% confidence if the pop. std. dev. is 0.88 mpg?

\[
\text{m.o.e.} = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}
\]

Example 3: A car rental agency wants to estimate the average mileage driven by its customers. A sample of 225 customer receipts, selected at random, yields an average of 325 miles. Assuming that the population standard deviation is 120 miles, construct a 95% confidence interval for the population average mileage.

Example 4: An insurance company collected data to estimate the mean value of personal property held by apartment renters in Tuscaloosa. In a random sample of 45 renters, the average value of personal property was $14,280 and the standard deviation was $6,540.
a. At 95% confidence, find the margin of error for estimating the true population mean?

Example 4:

b. Construct a 95% confidence interval for the true mean value of personal property owned by renters in Tuscaloosa.

Example 4:

c. Three years ago the true mean value of personal property owned by renters in Tuscaloosa was $13,050.

Based on your confidence interval, is there evidence that the true mean has changed?

Example 4:

d. Assuming that the true mean is unchanged from three years ago, find the probability that a future sample of 45 renters would result in a mean that is more extreme than the sample the insurance company just took.

\[ P(\bar{X} > 14,280 \mid \text{true mean} = 13,050) \]

What distribution must be used?

Example 4:

e. Based on this probability, would you conclude that the true mean has changed from three years ago?

or, equivalently

Is there sufficient evidence to conclude that the true mean has changed?

or, equivalently

Was the sample mean of $14,280 too close to $13,050 to call it unusual, or was this value a rare event?

Had the sample mean been $15,000, would your conclusion be different?