If you were asked to estimate someone’s height from memory, without that person being in front of you, you might say the person is about 5-11, 5 feet 11 inches. If a better estimate was needed, you might say between 5-10 and 6 feet. What would your confidence level be? Are you about 90% confident? If you needed to be more confident, say 95% confident, you might say between 5-8 and 6-1. The “interval” is larger to allow for less error. So your answer is your estimate plus or minus an estimating error or error of estimate. In this case, it’s 5-11 ± 2 inches or 5’11” ± 2”.

Now let’s look at a schematic of a population and a sample; \( \mu \) represent the population mean and \( \sigma \) represent the population standard deviation. These are values we may never know due to the size of the population, but they can be estimated from a sample, as long as the sample is random and representative of the population. So, if you want to estimate the population mean, you will need a “point estimate” of it. That would come from the sample mean, \( \bar{x} \). Your estimate is also dependent on the spread of the data. What measures the spread of the data? Isn’t that the standard deviation? Finally, you have to decide or be told how accurate your estimate needs to be. Do you want to be 90% accurate, 95% accurate, or 99% accurate? What “number” could be used to represent these levels of accuracy? It would be a “z” value; 90% accuracy can be represented by the number 1.645; 95% by 1.96 and 99% by 2.58.

Let’s call the error of estimate “E”; E must have the following components: A measure of spread - we’ll use the standard deviation; a sample size, \( n \), and a degree of accuracy - we’ll use a z score.

\[
E = z \frac{\sigma}{\sqrt{n}}
\]

Let’s try an example. If \( \bar{x} \) of a set of data for the amount of cereal in 35 boxes is 14.09 ounces and the standard deviation, \( \sigma \), is 0.008, find a 95% confidence interval for \( \mu \), the population mean for all the cereal boxes produced that day.

What is \( z \)? What is \( E \)? Add and subtract \( E \) from the point estimate of \( \mu \), the sample mean \( \bar{x} \).

\[
Z = \text{invNorm}(1 - .95/2) = \text{invNorm}(0.025) = -1.96, \text{ so } z = 1.96; \quad E = 1.96(0.008)/\sqrt{35} \approx 0.0027\text{ or, if rounded to three decimal places, 0.003.}
\]

\[
14.09 - 0.003 < \mu < 14.09 + 0.003 \quad \text{or} \quad 14.087 \leq \mu \leq 14.093
\]

Wow, I hope this can be done on the calculator! It can be! Look at the next page.
Focus – Estimating the Population Mean

Press STAT, scroll over to TESTS and choose #7: Z Interval. Choose stats, enter sigma = 0.008, the sample mean = 14.09 and n = 35; choose the confidence level = 95 and “Calculate”; press ENTER. There’s your interval: 14.087 ≤ µ ≤ 14.093 To find E, take the upper limit and subtract the mean: 14.093 – 14.09 = 0.003.

If you have raw data, put the data in your list menu and choose “Data” for your Z-Interval. Put in the appropriate information to get your interval.

t-distribution:

The above process assumes you know the population standard deviation, σ. How realistic is that? If you are trying to estimate the population mean, µ, from a sample, in most cases, how would you ever know the population standard deviation? You wouldn’t know it!

Since you don’t know the population standard deviation, σ, you’ll have to use the sample standard deviation S_x. You will be less confident in your estimate. If your data are approximately normal with no outliers as indicated from the box-and-whisker plot and normal probability plot, you may use the t-distribution. The t-distribution is a flattened version of the normal distribution – bell-shaped and symmetric. The t values for the confidence intervals are larger than the z values for the same confidence levels; the area under the curve is still 1 or 100% but since the distribution is wider, the t values are larger than “z”. The table of t-values also depends on the sample size, so the tables are actually a series of distributions, as your instructor will tell you. The t-distribution you use depends on the “degrees of freedom” which is simply d.f. = n – 1. Doing this all on the calculator, you are simply going to choose #8 under “TESTS”, the T-Interval. Here is an example problem.

Data were gathered on the response minutes for answering the phone at an 800 number. The sample mean, \( \bar{x} \), time was 6.72 minutes and the sample standard deviation, S_x, was 1.43 minutes from a sample size, n, of 74 data samples. Find a 95% confidence interval for µ, the population mean for the data for that day of the week and that time of day. E = t \( \frac{S_x}{\sqrt{n}} \); using the tables, t for n = 74 and confidence level = 95% is 1.667. Here’s how you get that; n = 74 so degrees of freedom = 73; there is no row 73, so drop back to row 70. The column is “0.05” or 1 – .95. E = 1.667 (1.43) / \sqrt{74} \approx 0.277 or 0.28 minutes. 6.72 – 0.28 ≤ µ ≤ 6.72 + 0.28 which is 6.44 ≤ µ ≤ 7.00 minutes.

Using the calculator, you will get a different result since the calculator contains the distribution for 73 degrees of freedom. Go to “STAT”, “TESTS” and choose T-Interval; put in the sample mean, sample standard deviation, sample size and confidence level. Notice it takes a longer time to get an answer since there are more calculations to do. To get E, take the upper bound of the interval and subtract the mean. What did you get? __________ I got \( \approx \) (about) 0.33 minutes (0.3313). The answer is 6.39 ≤ µ ≤ 7.05 minutes.

Let’s use some raw data and do a problem. A sample of the weight of 15 bags of sugar were examined from a bagging machine during a one-hour period. Adjustments were made for the weight of the packaging. The purpose was to check to see if the machine is bagging the sugar at the proper weight. Use the weights to get a 95% confidence interval for the population mean weight for this machine for this day of production.

<table>
<thead>
<tr>
<th>5.04</th>
<th>5.02</th>
<th>5.03</th>
<th>5.05</th>
<th>5.04</th>
<th>5.03</th>
<th>5.05</th>
<th>5.05</th>
<th>5.05</th>
<th>5.06</th>
<th>5.08</th>
<th>5.03</th>
<th>5.04</th>
<th>5.03</th>
</tr>
</thead>
</table>

Find the sample mean and standard deviation. Use the t-tables to find the t value for n = 15 and level = 95%. Find E, the estimate of error showing the set-up for it. Find the interval. Now use the calculator to find the same interval. What is E?

\( \bar{x} = 5.04 \) pounds; \( S_x = 0.016 \) pounds; \( t = 1.761 \) (row 14 for d.f. and column 0.05); E = 1.761 (0.016) / \sqrt{15} \approx 0.007; the “long-hand” interval is 5.04 – 0.007 ≤ µ ≤ 5.04 + 0.007 ; 5.033 ≤ µ ≤ 5.047; calculator answer: 5.031 ≤ µ ≤ 5.049; E = 0.0089 = upper bound of interval – mean.

HANDOUTS at www.hacc.edu Choose “Lancaster Campus” from the “pull-down” menu.

Scroll to bottom and choose “Tutoring” under the “STUDENTS” heading. Choose Lancaster Tutoring” on the left side. Scroll down to “Guides for TI 83 or 84 graphing calculators” – in red letters – click on it. Scroll down to “Handouts”. If you are unfamiliar with the calculator, choose
“Calculator Handout – Basic Operations” first or choose the topic you are learning and if you need to do so, print it. You can also come to the Lancaster Tutoring Center for free personal help!