Confidence Intervals for the Mean (σ unknown)

In many real life situations, the standard deviation is unknown. In order to construct a confidence interval for a random variable that is normally distributed and the standard deviation is unknown, you can use a t-distribution.

If the distribution of a random variable $x$ is approximately normal, then

$$ t = \frac{\bar{x} - \mu}{s / \sqrt{n}} $$

follows a t-distribution.

Critical values of $t$ are denoted by $t_c$. Here are several properties of the t-distribution:

1. The mean, median and mode of the t-distribution are equal to 0.
2. The t-distribution is bell-shaped and symmetric around the mean.
3. The total area under the t-distribution is equal to 1.
4. The tails in the t-distribution are "thicker" than those of a standard normal distribution.
5. The standard deviation of the t-distribution varies with the sample size, but it is greater than 1.
6. The t-distribution is a family of curves, each determined by a parameter called the degrees of freedom. The degrees of freedom (sometimes abbreviated d.f.) are the number of free choices left after a sample statistic such as $x$ is calculated. When you use a t-distribution to estimate a population mean, the degrees of freedom are equal to one less than the sample size.

$$ d.f. = n - 1 $$

7. As the degrees of freedom increase, the t-distribution approaches the standard normal distribution, as shown in the figure. After 30 d.f., the t-distribution is close to the standard normal distribution.
Finding the Critical Values of t

Find the critical value $t_c$ for a 95% confidence interval when the sample size is 15.

Because $n = 15$, the degrees of freedom are

$$d.f. = n - 1 = 15 - 1 = 14$$

Using Table 5 in Appendix B lists the critical values for $t$ for selected confidence intervals and degrees of freedom. Find the confidence level you are given, then go down to the given degrees of freedom.

0.95 level of confidence with 14 degrees of freedom is 2.145.

Interpretation: For a $t$ - distribution curve with 14 degrees of freedom, 95% of the curves lies between $t = \pm 2.45$.

TRY: Find the critical value $t_c$ for a 90% confidence level when the sample size is 22.

When the degrees of freedom you need is not on the table, use the closest d.f. in the table that is less than the value that you need. For instance, for d.f. = 57, use 50 degrees of freedom.
Confidence Intervals and t-Distributions

Constructing a confidence interval for $\mu$ when $\sigma$ is not known using the t-distribution is similar to constructing a confidence interval for $\mu$ when $\sigma$ is known using the standard normal distribution.

**Constructing a Confidence Interval for a Population Mean ($\sigma$ unknown)**

1. Verify that $\sigma$ is not known, the sample is random, and either the population is normally distributed or $n \geq 30$.
2. Find the sample statistics $n$, $x$, and $s$.
3. Identify the degrees of freedom, the level of confidence $c$, and the critical value $t_c$.
4. Find the margin of error $E$.
5. Find the left and right endpoints and form the confidence interval.

\[
x = \frac{\sum x}{n}, \quad s = \sqrt{\frac{\sum (x - \bar{x})}{n - 1}}
\]

\[d.f. = n - 1\]

Use Table 5 in Appendix B.

\[E = t_c \frac{s}{\sqrt{n}}\]

Left endpoint: $x - E$

Right endpoint: $x + E$

Interval: $x - E < \mu < x + E$
You randomly select 16 coffee shops and measure the temperature of the coffee sold at each. The sample mean temperature is 162.0°F with a sample standard deviation of 10.0°F. Construct a 95% confidence interval for the population mean temperature of coffee sold. Assume the temperatures are approximately normally distributed.

\[ n = 16 \quad \bar{x} = 162.0 \quad s = 10.0 \quad c = 0.95 \quad d.f. = 16 - 1 = 15 \]

Find \( t_c \) using the table. 15 d.f. with a .95 confidence level. \( t_c = 2.131 \)

Now find \( E \).

\[ E = t_c \cdot \frac{s}{\sqrt{n}} = 2.131 \cdot \frac{10.0}{\sqrt{16}} \approx 5.3 \]

Find the endpoints

Left Endpoint  \hspace{1cm} Right Endpoint

\[ \bar{x} - E = 162.0 - 5.3 = 156.7 \quad \bar{x} + E = 162.0 + 5.3 = 167.3 \]

**INTERVAL**

\[ 156.7 < \mu < 167.3 \]
Using the Calculator

STAT TESTS scroll down to 8 (TInterval) Inpt: Stats

\( \overline{x} : 162.0 \quad Sx : 10.0 \quad n : 16 \quad C - Level : 0.95 \)

TRY:

Construct 90% and a 99% confidence intervals for the population mean temperature of coffee sold using the previous example.
You randomly select 36 cars of the same model that were sold at a car dealership and determine the number of days each car sat on the dealership's lot before it was sold. The sample mean is 9.75 days, with a sample standard deviation of 2.39 days. Construct a 99% confidence interval for the population mean number of days the car model sits on the dealership's lot.

\[ n = 36 \quad \bar{x} = 9.75 \quad s = 2.39 \quad c = 0.99 \quad d.f. = 35 \]

using the table:

\[ t_c = 2.724 \]

\[ E = t_c \frac{s}{\sqrt{n}} = 2.724 \times \frac{2.39}{\sqrt{36}} \approx 1.09 \]

Left Endpoint  
\[ \bar{x} - E = 9.75 - 1.09 = 8.66 \]

Right Endpoint  
\[ \bar{x} + E = 9.75 + 1.09 = 10.84 \]

Confidence interval  
\[ 8.66 < \mu < 10.84 \]

Calculator:
STAT  TESTS  8) TInterval  
\[ \bar{x} = 9.75 \quad Sx = 2.39 \quad n = 36 \quad c - Level = 0.99 \]

Try the same problem but with a 90% confidence level.
Choosing the Standard Normal Distribution or the \( t \)-Distribution

The flow chart describes when to use the standard normal distribution and when to use the \( t \)-distribution to construct a confidence interval for a population mean.

\[
E = z_c \frac{\sigma}{\sqrt{n}} \quad \text{Section 6.1}
\]

\[
E = t_c \frac{s}{\sqrt{n}} \quad \text{Section 6.2}
\]

and \( n - 1 \) degrees of freedom.

Notice in the flowchart that when \( n < 30 \) and the population is not normally distributed, you cannot use the standard normal distribution or the \( t \)-distribution.
You randomly select 25 newly constructed houses. The sample mean of the construction cost is $181,000 and the population standard deviation is $28,000. Assuming construction costs are normally distributed, should you use the standard normal distribution, the t-distribution, or neither to construct a 95% confidence interval for the population mean construction cost.

Is $\sigma$ known?  
Yes

Is either the population normally distributed or $n \geq 30$?  
Yes

Decision:  
Use the standard normal distribution.

TRY:  
You randomly select 18 adult male athletes and measure the resting heart rate of each. The sample mean heart rate is 64 beats per minute, with a sample standard deviation of 2.5 beats per minute. Assuming the heart rates are normally distributed, should you use the standard normal distribution, the t-distribution, or neither to construct a 90% confidence interval for the population mean heart rate? Explain.

Is $\sigma$ known?  

Is either the population normally distributed or $n \geq 30$?  

Decision:
P. 315  1 - 35 odds
(#29 just construct the confidence interval)