Section 7-3, 7.4
Estimating a Population Mean

Assumptions:
1. The sample is a simple random sample.
2. The population is normally distributed or \( n > 30 \).

---

Point estimate

- Population mean: \( \mu \) (unknown)
- Point Estimate:
  - The sample mean: \( \bar{x} \)
  - Exact standard error: \( \frac{\sigma}{\sqrt{n}} \)
  - Estimated standard error (se): \( \frac{\bar{x}}{\sqrt{n}} \)

---

(i) 100(1-\(\alpha\))% Confidence Interval for Population Mean \( \mu \) if \( \sigma \) is known

Point estimate \( \pm \) margin of error

\[ \bar{x} \pm E \]

\[ (\bar{x} - E, \bar{x} + E) \]

\[ E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \]

---

Example: Constructing confidence interval

A study found the body temperatures of 106 healthy adults. The sample mean was 98.2 degrees and the standard deviation was 0.62 degrees. Find

(a) the point estimate of the population mean \( \mu \) of all body temperatures. 98.2 degrees

(b) the margin of error \( E \)

(c) the 95% confidence interval for \( \mu \).

---

Solution: (a) 98.20

\[ n = 106 \]
\[ \bar{x} = 98.20 \]
\[ \sigma = 0.62 \]

\[ z_{0.05} = 1.96 \]

(b) \[ E = z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{0.62}{\sqrt{106}} = 0.12 \]

(c) Recall

\[ \bar{x} - E < \mu < \bar{x} + E \]

\[ 98.20 - 0.12 < \mu < 98.20 + 0.12 \]

\[ 98.08 < \mu < 98.32 \]

---

Sample Size for Estimating Mean \( \mu \)

\[ n = \frac{(z_{\alpha/2}) \cdot \sigma}{E}^2 \]

Sample size formula with 95% confidence level and margin of error \( E \) is approximated by if \( \sigma \) is unknown

\[ n = \frac{4s^2}{E^2} \]
**Round-Off Rule for Sample Size \( n \)**

When finding the sample size \( n \), if it does not result in a whole number, always increase the value of \( n \) to the next larger whole number.

**Finding the Sample Size \( n \) when \( \sigma \) is unknown**

1. Use the range rule of thumb to estimate the standard deviation as follows: \( \sigma = \text{range}/4 \).
2. Conduct a pilot study by starting the sampling process. Based on the first collection of at least 31 randomly selected sample values, calculate the sample standard deviation \( s \) and use it in place of \( \sigma \).
3. Estimate the value of \( \sigma \) by using the results of some other study that was done earlier.

**Example:**

Assume that we want to estimate the mean IQ score for the population of statistics professors. How many statistics professors must be randomly selected for IQ tests if we want 95% confidence that the sample mean is within 2 IQ points of the population mean? Assume that \( \sigma = 15 \), as is found in the general population.

\[
\begin{align*}
\alpha &= 0.05 \\
\alpha/2 &= 0.025 \\
2z_{0.025} &= 1.96 \\
E &= 2 \\
\sigma &= 15
\end{align*}
\]

With a simple random sample of only 217 statistics professors, we will be 95% confident that the sample mean will be within 2 points of the true population mean \( \mu \).

**Important Properties of the Student \( t \) Distribution**

1. The Student \( t \) distribution is different for different sample sizes (see Figure for the cases \( n = 3 \) and \( n = 12 \)).
2. The Student \( t \) distribution has the same general symmetric bell shape as the normal distribution but it reflects the greater variability (with wider distributions) that is expected with small samples.
3. The Student \( t \) distribution has a mean of \( t = 0 \) (just as the standard normal distribution has a mean of \( z = 0 \)).
4. The standard deviation of the Student \( t \) distribution varies with the sample size and is greater than 1 (unlike the standard normal distribution, which has a \( \sigma = 1 \)).
5. As the sample size \( n \) gets larger, the Student \( t \) distribution gets closer to the normal distribution.

**Table B**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>80%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
<th>99.9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( df )</td>
<td>( t_{0.10} )</td>
<td>( t_{0.05} )</td>
<td>( t_{0.025} )</td>
<td>( t_{0.01} )</td>
<td>( t_{0.001} )</td>
</tr>
<tr>
<td>1</td>
<td>3.078</td>
<td>6.314</td>
<td>12.706</td>
<td>31.821</td>
<td>63.657</td>
</tr>
<tr>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>6</td>
<td>1.640</td>
<td>1.943</td>
<td>2.682</td>
<td>3.143</td>
<td>3.708</td>
</tr>
<tr>
<td>7</td>
<td>1.415</td>
<td>1.895</td>
<td>2.365</td>
<td>2.990</td>
<td>3.499</td>
</tr>
</tbody>
</table>
If the distribution of a population is essentially normal, then the distribution of
\[ t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \]
\[ \text{is essentially a Student } t \text{ Distribution for all samples of size } n \]
\[ \text{Degrees of Freedom (df) = } n - 1 \]

**Margin of Error E for Estimating \( \mu \)**

Based on an Unknown \( \sigma \) and a Small Simple Random Sample from a Normally Distributed Population

\[ E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \]

where \( t_{\alpha/2} \) has \( n - 1 \) degrees of freedom.

\[ \cdot 100(1-\alpha)\% \text{Confidence Interval for } \mu \]

\[ \bar{x} \pm t_{\alpha/2} \left( \frac{s}{\sqrt{n}} \right); \text{ df } = n - 1 \]

- \( t_{\alpha/2} \) found in Table B
- Based on an Unknown \( \sigma \) and a Small Simple Random Sample from a Normally Distributed Population

\[ \cdot \text{A 95\% confidence interval for the population mean } \mu \text{ is:} \]

\[ \bar{x} \pm t_{0.025} \left( \frac{s}{\sqrt{n}} \right); \text{ df } = n - 1 \]

---

**The Standard Normal Distribution is the \( t \)-Distribution with df = \( \infty \)**

**Example:**

A study found the body temperatures of 106 healthy adults. The sample mean was 98.2 degrees and the sample standard deviation was 0.62 degrees. Find the margin of error \( E \) and the 95% confidence interval for \( \mu \).

\[ n = 106 \quad E = t_{0.025} \cdot \frac{s}{\sqrt{n}} = 1.984 \cdot 0.62 \sqrt{106} = 0.62 \]

\[ x = 98.20^\circ \quad s = 0.62^\circ \]

\[ \alpha = 0.05 \quad \alpha/2 = 0.025 \quad 98.20^\circ - 0.62 \cdot 0.025 < \mu < 98.20^\circ + 0.62 \cdot 0.025 \]

\[ t_{0.025} = 1.984 \]

The interval is the same here as in Section 6-2, but in some other cases, the difference would be much greater.
Summary: Sections 7.1-7.4

- Point estimates:
  \[ \hat{p} \quad \text{or} \quad \bar{x} \to p \quad \text{or} \quad \mu \]

- Margin of Error: \( E = (\text{critical value})(\text{standard error}) \)

- Confidence Interval: \( (\text{point estimate}) \pm E \)

- Sample size \( n \): a solution of \( E = (\text{critical value})(\text{standard error}) \)