Econ 3790: Business and Economics Statistics

Instructor: Yogesh Uppal
Email: yuppal@ysu.edu
Chapter 8: Interval Estimation

- Population Mean: $\sigma$ Known
- Population Mean: $\sigma$ Unknown
A point estimator cannot be expected to provide the exact value of the population parameter.

An interval estimate can be computed by adding and subtracting a margin of error to the point estimate.

Point Estimate $+/-$ Margin of Error

The purpose of an interval estimate is to provide information about how close the point estimate is to the value of the parameter.
The general form of an interval estimate of a population mean is

\[ \bar{x} \pm \text{Margin of Error} \]

In order to develop an interval estimate of a population mean, the margin of error must be computed using either:

- the population standard deviation \( \sigma \), or
- the sample standard deviation \( s \)

These are also \textit{Confidence Interval}. 
Interval Estimate of a Population Mean: $\sigma$ Known

- Interval Estimate of $\mu$

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

where:
- $\bar{x}$ is the sample mean
- $1 - \alpha$ is the confidence coefficient
- $z_{\alpha/2}$ is the $z$ value providing an area of $\alpha/2$ in the upper tail of the standard normal probability distribution
- $\sigma$ is the population standard deviation
- $n$ is the sample size
Interval Estimation of a Population Mean: \( \sigma \) Known

- There is a \( 1 - \alpha \) probability that the value of a sample mean will provide a margin of error of \( z_{\alpha/2} \sigma_x \) or less.
Summary of Point Estimates Obtained from a Simple Random Sample

<table>
<thead>
<tr>
<th>Population Parameter</th>
<th>Parameter Value</th>
<th>Point Estimator</th>
<th>Point Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$ = Population mean</td>
<td>40.9</td>
<td>$\bar{x}$ = Sample mean</td>
<td></td>
</tr>
<tr>
<td>$\sigma$ = Population std. deviation</td>
<td>20.5</td>
<td>$s$ = Sample std. deviation</td>
<td>$\ldots\ldots$</td>
</tr>
<tr>
<td>$p$ = Population proportion</td>
<td>.62</td>
<td>$\bar{p}$ = Sample proportion</td>
<td></td>
</tr>
</tbody>
</table>
Example: Air Quality

- Consider our air quality example. Suppose the population is *approximately normal* with $\mu = 40.9$ and $\sigma = 20.5$. This is *$\sigma$ known* case.
- If you guys remember, we picked a sample of size 5 ($n = 5$).
- Given all this information, What is the margin of error at 95% confidence level?
Example: Air Quality

- What is the margin of error at 95% confidence level.
  \[ z_{0.025} \frac{\sigma}{\sqrt{n}} = 1.96 \times \frac{20.5}{\sqrt{5}} = 1.96 \times 9.2 = 18 \]

- We can say with 95% confidence that population mean (\( \mu \)) is between \( \pm 18 \) of the sample mean.
  \[ \bar{x} \pm z_{0.025} \frac{\sigma}{\sqrt{n}} \]

- With 95% confidence, \( \mu \) is between …. and …..
Interval Estimation of a Population Mean: $\sigma$ Unknown

- If an estimate of the population standard deviation $\sigma$ cannot be developed prior to sampling, we use the sample standard deviation $s$ to estimate $\sigma$.
- This is the $\sigma$ unknown case.
- In this case, the interval estimate for $\mu$ is based on the $t$ distribution.
- (We’ll assume for now that the population is normally distributed.)
The *t* distribution is a family of similar probability distributions.

A specific *t* distribution depends on a parameter known as the degrees of freedom.

Degrees of freedom refer to the number of independent pieces of information that go into the computation of *s*. 
$t$ Distribution

A $t$ distribution with more degrees of freedom has less dispersion.

As the number of degrees of freedom increases, the difference between the $t$ distribution and the standard normal probability distribution becomes smaller and smaller.
t Distribution

Standard normal distribution

t distribution (20 degrees of freedom)

t distribution (10 degrees of freedom)
For more than 100 degrees of freedom, the standard normal $z$ value provides a good approximation to the $t$ value.

The standard normal $z$ values can be found in the infinite degrees ($\infty$) row of the $t$ distribution table.
### t Distribution

<table>
<thead>
<tr>
<th>Degrees of Freedom</th>
<th>Area in Upper Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>50</td>
<td>0.849</td>
</tr>
<tr>
<td>60</td>
<td>0.848</td>
</tr>
<tr>
<td>80</td>
<td>0.846</td>
</tr>
<tr>
<td>100</td>
<td>0.845</td>
</tr>
<tr>
<td>∞</td>
<td>0.842</td>
</tr>
</tbody>
</table>

**Standard normal z values**
Interval Estimation of a Population Mean: 
\( \sigma \) Unknown

- Interval Estimate

\[
\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}
\]

where: 
- \( 1 - \alpha \) = the confidence coefficient
- \( t_{\alpha/2} \) = the \( t \) value providing an area of \( \alpha / 2 \) in the upper tail of a \( t \) distribution with \( n - 1 \) degrees of freedom
- \( s \) = the sample standard deviation
Example: Air quality when $\sigma$ is unknown

- Now suppose that you did not know what $\sigma$ is. You can estimate using the sample and then use t-distribution to find the margin of error.

- What is 95% confidence interval in this case? The sample size $n = 5$. So, the degrees of freedom for the t-distribution is 4. The level of significance ($\alpha$) is 0.05. $s = \ldots$

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$
Summary of Interval Estimation Procedures for a Population Mean

Can the population standard deviation $\sigma$ be assumed known?

Yes

$\sigma$ Known Case

Use

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

No

Use the sample standard deviation $s$ to estimate $\sigma$

$\sigma$ Unknown Case

Use

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$
Interval Estimation of a Population Proportion

The general form of an interval estimate of a population proportion is

\[ \hat{p} \pm \text{Margin of Error} \]
Interval Estimation of a Population Proportion

- Interval Estimate

\[
\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}
\]

where:

- \(1 - \alpha\) is the confidence coefficient
- \(z_{\alpha/2}\) is the \(z\) value providing an area of \(\alpha/2\) in the upper tail of the standard normal probability distribution
- \(\bar{p}\) is the sample proportion