Warm-up Remarks:

- This is the first introduction to one of the two major activities of inferential statistics.
  - Thus far we have used descriptive statistics to summarize or describe important characteristics of data, but with inferential statistics we use sample data to make inferences (or generalizations) about a population.
- The two major applications of inferential statistics involve the use of sample data to:
  1. estimate the value of a population parameter
  2. test some claim (or hypothesis) about a population

Today we will study:

- How to find a point estimate of a population parameter
- How to find a margin of error, given a level of confidence
- How to construct and interpret confidence intervals for the population mean

---

**Estimating Population Parameters**

Paraphrased from our text - section 6.1

The requirements for the methods of this section:

1. The sample size is at least 30 or the population is normally distributed and the population standard deviation $\sigma$ is known.
2. The sample is a simple random sample.

**DEFINITION**

A **point estimate** is a single value used to approximate a population parameter. The sample mean $\bar{x}$ is the best point estimate of the population mean $\mu$.

We have no estimate of how good our estimate is - the probability is essentially zero that the estimate is exactly correct!

In order to reveal how good the estimate is we use a:

**DEFINITION**

A **confidence interval** (or **interval estimate**) is an interval, or a range of values used to estimate the true value of a population parameter. A confidence interval is often abbreviated as **CI**.

A confidence interval is associated with a degree of confidence (say 95% or 0.95).

**DEFINITION**

A **level of confidence**, denoted by $c$, is the proportion of times that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times.

**Constructing a Confidence Interval**

**DEFINITION**

A **critical value** $z_c$ is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely to appear.

$z_c$ is the positive $z$-score that is at the vertical boundary separating an area of $\frac{1}{2}(1 - c)$ in the right tail of the standard normal distribution (the value of $-z_c$ is at the vertical boundary for the area of $\frac{1}{2}(1 - c)$ in the left tail.
Remarks

1. The central limit theorem tells us that when \( n \geq 30 \) the sampling distribution of sample means is a normal distribution.

2. The level of confidence \( c \) is the area under the standard normal curve between the critical values \(-z_c\) and \(z_c\).

3. The area in the tails (outside \( c \)) is \(1 - c\), so the area in each tail is \(\frac{1}{2}(1 - c)\).

The difference between the point estimate and the actual population parameter is called the sampling error. In the case of the mean this would reflect the difference \(\bar{x} - \mu\).

However, \( \mu \) is unknown (or why would we wish to estimate it) and \(\bar{x}\) will vary from sample to sample.

**DEFINITION**

Given a level of confidence \( c \), the margin of error, denoted by \( E \), is the maximum likely difference (with probability \( c \)) between the sample point estimate and the true value of the population parameter it is estimating. The margin of error is sometimes called the maximum error of the estimate and can be found by multiplying the critical value and the standard deviation of sample means.

\[
E = z_c \sigma_{\bar{x}} = z_c \frac{\sigma}{\sqrt{n}}
\]

When \( n \geq 30 \), the sample standard deviation \( s \) can be used in place of \( \sigma \).

Confidence Intervals for the Population Mean

**DEFINITION**

A \( c \)-confidence interval for the population mean \( \mu \) is

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

The probability that the confidence interval contains \( \mu \) is \( c \). Often written in interval notation: \((\bar{x} - E, \bar{x} + E)\).

Round-off Rule

1. Round to the same number of decimal places given for the sample mean.
2. When using the original sample data - round to one more decimal place than used for the original data.
3. Do not round until the end (keep all the decimal places during intermediate steps).

**GUIDELINES**

**Constructing a Confidence Interval for the Mean**

\((n \geq 30 \text{ or } \sigma \text{ known; with a normally distributed population)}\)

1. Find the sample statistics \( n \) and \(\bar{x}\).
2. Specify \( \sigma \), if known. Otherwise, if \( n \geq 30 \), find the sample standard deviation \( s \) and use is as an estimate for \( \sigma \).
3. Find the critical value \( z_c \) that corresponds to the given level of confidence.
4. Find the margin of error \( E \).

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

5. Find the left and right endpoints and write down the confidence interval.

\[
\text{Interval: } \bar{x} - E < \mu < \bar{x} + E
\]
You must be careful when you interpret the meaning of a confidence interval.

- **Class Example (95% CI)** -

<table>
<thead>
<tr>
<th>Interpreting a Confidence Interval</th>
<th>Given: $98.08 &lt; \mu &lt; 98.32$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correct:</strong> We are 95% confident that the interval 98.08 to 98.32 actually does contain the true value of $\mu$.</td>
<td></td>
</tr>
<tr>
<td><strong>Wrong:</strong> There is a 95% chance that the true value of $\mu$ will fall between 98.08 and 98.32.</td>
<td></td>
</tr>
</tbody>
</table>

**Remarks**
- The population mean $\mu$ exists and is a fixed value, either the limits enclose the value $\mu$ or they do not.