7.4 Estimating a Population Mean: σ Not Known

In section 7.3, we constructed confidence intervals for the mean of a population whose standard deviation σ was known. This assumption is not very realistic. In this section we consider populations where σ is not known. We collect sample data and compute n, $\overline{x}$ *and* s. The methods of this section are realistic and practical and are used often. They are based on two assumptions which do not include the need to know σ:

ASSUMPTIONS

1. The sample is a simple random sample
2. Either the sample is from a normally distributed population or $n > 30$.

POINT ESTIMATE of μ in these cases is still $\overline{x}$.

Now if the above two conditions are satisfied, to set up the **CONFIDENCE INTERVAL** estimates of μ, instead of using the normal distribution, we use the Student t distribution:

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

In a student t Distribution, **CRITICAL VALUES** $t_{\alpha/2}$ are found in table A-3 by locating the degrees of freedom and the value of α for the two tails.

DEGREES OF FREEDOM: for a single data set is the number of sample values that can vary after certain restrictions have been imposed on all data values.

degrees of freedom = $n - 1$ (one less than the sample size)

READING THE STUDENT t TABLES: (Two-tailed)
Suppose the conditions are satisfied to use the t table (table A-3). Find the critical value $t_{\alpha/2}$ for the given sample size and degree of confidence:

(a) $n = 20$, 95%

(b) $n = 27$, 90%

(c) $n = 16$, 99%

PROPERTIES OF THE STUDENT t DISTRIBUTION

1. Different for different sample sizes. See figure 7-5 on page 359
2. Same general symmetric bell shape as the standard normal distribution, but with wider distributions. (bell curves are flatter)
3. Has mean of $t = 0$.
4. Standard deviation varies with the sample size, but it is greater than 1.
5. As n gets larger, the Student t distribution gets closer to the standard normal distribution.

CONDITIONS FOR USING THE STUDENT T DISTRIBUTION

1. σ is unknown, (if σ is known we use the methods of 6.3);
2. The parent population has a distribution that is essentially normal; or
3. If the parent population is not normally distributed, then $n > 30$.

Table 6-1	**Choosing Between z and t**
Method:	Conditions:
Use Normal (z) Distribution	σ is known and normally distributed population, or
	σ is known and $n > 30$
Use t distribution	σ is not known and normally distributed population, or
	σ is not known and $n > 30$
Use methods of Chapter 12	Population is not normally distributed and $n \leq 30$

Notes: 1. **Criteria for deciding whether the population is normally distributed**: Population need not be exactly normal, but it should appear to be somewhat symmetric with one mode and no outliers.

2. **Sample size $n > 30$:** This is a commonly used guideline, but sample sizes of 15 to 30 are adequate if the population appears to have a distribution that is notfar from being normal and there are no outliers. For some population distributions that ae extremely far from normal, the sample size might need to be larger than 50 or even 100.

CHOOSE THE APPROPRIATE DISTRIBUTION z Versus t

***** #6, pg. 365**

***** #8, pg. 365**

*****#10, pg. 365**

MARGIN OF ERROR (will tend to be larger for student t)

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$ where $t_{\alpha/2}$ has n-1 degrees of freedom

CONFIDENCE INTERVAL: $\bar{x} - E < \mu < \bar{x} + E$

where $E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$

*****#14, pg. 366**

*****#16, pg. 366 (Use the calculator screen display provided)**

*****#18, pg. 366**

*****#20, pg. 367**

Using the TI-83: This time under STAT>>TESTS, choose #8 (TInterval)

*****#20, pg. 367**

RECAP OF CHAPTER 7

SECTION 7.2 **Estimating a Population Proportion (round to 3 significant digits)**

$\widehat{\mathbf{p}} = \frac{x}{n}$ is best point estimate of a population proportion

STANDARD DEVIATION OF SAMPLE PROPORTIONS: $\sigma_{\text{sample prop}} = \sqrt{\frac{\widehat{p}\widehat{q}}{n}}$

MARGIN OF ERROR OF THE ESTIMATE OF p: $E = z_{\alpha/2}\sqrt{\frac{\widehat{p}\widehat{q}}{n}}$

CONFIDENCE INTERVAL FOR THE POPULATION PROPORTION p:

$$\widehat{p} - E < p < \widehat{p} + E$$

DETERMINING SAMPLE SIZE (always round up)

If we have an estimate of p: $n = \frac{[z_{\alpha/2}]^2\widehat{p}\widehat{q}}{E^2}$

If no estimate of p is known: $n = \frac{[z_{\alpha/2}]^2 0.25}{E^2}$

SECTION 7.3 **Estimating a Population Mean: σ Known**

(If data is original round to one more decimal place than original values
If summary data is given, round to same accuracy as $\overline{x}$)

$\overline{x}$ is the best point estimate of the population mean μ

Use CRITICAL VALUES: $\pm z_{\alpha/2}$ **(table A-2)**

MARGIN OF ERROR: $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

CONFIDENCE INTERVAL: $\overline{x} - E < \mu < \overline{x} + E$

Determining Sample Size (always round up): $n = [\frac{z_{\alpha/2} \cdot \sigma}{E}]^2$**,**
where $z_{\alpha/2}$ is the critical z score based on the desired degree of confidence, E is the desired margin of error, and σ is the population standard deviation.

SECTION 7.4 **Estimating a Population Mean: σ Unknown**

When σ is unknown and the population is normally distributed, or σ is unknown and n $\leq$ 30:

Use **Student t-scores** rather than z-scores **(table A-3)**

MARGIN OF ERROR (will tend to be larger for student t)

$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$ where $t_{\alpha/2}$ has n-1 degrees of freedom

CONFIDENCE INTERVAL: $\overline{x} - E < \mu < \overline{x} + E$