Section 7-2

Estimating a Population Mean

SAMPLE MEANS

1. For many populations, the distribution of sample means $\bar{x}$ tends to be more consistent (with less variation) than the distributions of other sample statistics.
2. For all populations, the sample mean $\bar{x}$ is an unbiased estimator of the population mean μ, meaning that the distribution of sample means tends to center about the value of the population mean μ.

POINT ESTIMATE

Point Estimate: The sample mean $\bar{x}$ is the best *point estimate* (or single value estimate) of the population mean μ.

COMMENT

It is rare that we want to estimate the unknown value of a population mean but we somehow know the value of the population standard deviation σ. The realistic situation is that σ is not known. (We begin this section by considering this more realistic scenario.) When σ is not known, we construct the confidence interval by using the Student t distribution instead of the standard normal distribution.

ASSUMPTIONS FOR CONFIDENCE INTERVAL OF MEAN WITH σ NOT KNOWN

1. The sample is a simple random sample.
2. Either or both of the following conditions are satisfied:
   - The population is normally distributed
   - $n > 30$

THE STUDENT *t* DISTRIBUTION

If a population has a normal distribution, then the distribution of

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

is a **Student *t* distribution** for all samples of size n. The Student t distribution is often referred to as the ***t* distribution**.

DEGREES OF FREEDOM

Finding a critical value $t_{\alpha/2}$ requires a value for the **degrees of freedom** (or **df**). In general, the number of degrees of freedom for a collection of sample data is the number of sample values that vary after certain restraints have been imposed on the data values. For the methods of this section, the number of degrees of freedom is the sample size minus 1; that is,

$$\text{degrees of freedom} = n - 1$$

FINDING THE CRITICAL VALUE

A critical value $t_{\alpha/2}$ can be found using Table A-3 which is found on page 586, inside the back cover, and on the *Formulas and Tables* card. If the table does not include the number of degrees of freedom that you need, you could

- use the closest value
- be conservative and using the next lower number of degrees of freedom
- interpolate. For example, if you have 55 degrees of freedom, you could find the mean of the critical values for 50 and 60.

To keep things simple, ***we will use the closest value***.

MARGIN OF ERROR ESTIMATE OF μ (WITH σ NOT KNOWN)

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

where $(1-\alpha)$ is the confidence level and $t_{\alpha/2}$ has $n-1$ degrees of freedom.

CONFIDENCE INTERVAL ESTIMATE OF THE POPULATION MEAN μ (WITH σ NOT KNOWN)

$$\bar{x} - E < \mu < \bar{x} + E$$

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

CONSTRUCTING A CONFIDENCE INTERVAL FOR μ (σ NOT KNOWN)

1. Verify that the two required assumptions are met.
2. With σ unknown (as is usually the case), use $n - 1$ degrees of freedom and refer to Table A-3 to find the critical value $t_{\alpha/2}$ that corresponds to the desired confidence interval. (For the confidence level, refer to "Area in Two Tails.")
3. Evaluate the margin of error $E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$
4. Find the values of $\bar{x} - E$ and $\bar{x} + E$. Substitute these in the general format of the confidence interval: $\bar{x} - E < \mu < \bar{x} + E$.
5. Round the result using the same round-off rule on the following slide.

ROUND-OFF RULE FOR CONFIDENCE INTERVALS USED TO ESTIMATE μ

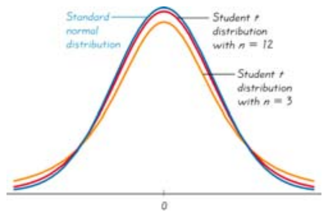
1. When using the ***original set of data*** to construct the confidence interval, round the confidence interval limits to ***one more decimal place*** than is used for the original data set.
2. When the original set of data is unknown and only the ***summary statistics*** $(n, \bar{x}, s)$ are used, round the confidence interval limits to the ***same number of places*** as used for the sample mean.

FINDING A CONFIDENCE INTERVAL FOR μ WITH TI-83/84

1. Select **STAT**.
2. Arrow right to **TESTS**.
3. Select 8**:TInterval….**
4. Select input (**Inpt**) type: **Data** or **Stats**. (Most of the time we will use **Stats**.)
5. Enter the sample mean, $\bar{x}$.
6. Enter the sample standard deviation, **Sx**.
7. Enter the size of the sample, **n**.
8. Enter the confidence level (**C-Level**).
9. Arrow down to **Calculate** and press **ENTER**.

PROPERTIES OF THE STUDENT *t* DISTRIBUTION

1. The Student t distribution is different for different sample sizes (see Figure below for the cases $n = 3$ and $n = 12$).



PROPERTIES OF THE STUDENT *t* DISTRIBUTION (CONTINUED)

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.

ESTIMATING A MEAN WHEN σ IS KNOWN

Requirements:

1. The sample is a simple random sample.
2. Either or both of these conditions are satisfied: The population is normally distributed or $n > 30$.

Confidence Interval:

$$\bar{x} - E < \mu < \bar{x} + E$$

where the margin of error E is found from the following:

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

Note: The critical value $z_{\alpha/2}$ is found from Table A-2 (the standard normal distribution).

FINDING A CONFIDENCE INTERVAL FOR μ WITH TI-83/84

1. Select **STAT**.
2. Arrow right to **TESTS**.
3. Select **7:ZInterval….**
4. Select input (**Inpt**) type: **Data** or **Stats**. (Most of the time we will use **Stats**.)
5. Enter the standard deviation, **σ**.
6. Enter the sample mean, **$\bar{x}$**.
7. Enter the size of the sample, **n**.
8. Enter the confidence level (**C-Level**).
9. Arrow down to **Calculate** and press **ENTER**.

SAMPLE SIZE FOR ESTIMATING μ

$$n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E}\right]^2$$

where $z_{\alpha/2}$ = critical z score based on desired confidence level

E = desired margin of error

σ = population standard deviation

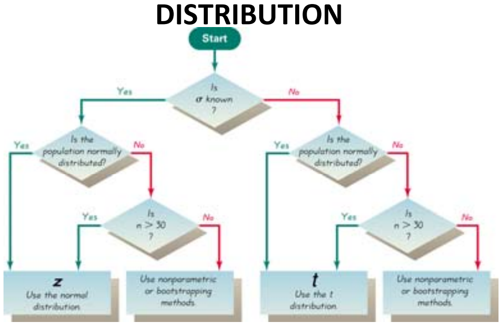
ROUND-OFF RULE FOR SAMPLE SIZE *n*

When finding the sample size n, if the use of the formula on the previous slide does not result in a whole number, always ***increase*** the value of n to the ***next larger*** whole number.

FINDING THE SAMPLE SIZE WHEN σ IS UNKNOWN

1. Use the range rule of thumb (see Section 3-3) to estimate the standard deviation as follows: $\sigma \approx$ range/4.
2. Start the sample collection process without knowing σ and, using the first several values, calculate the sample standard deviation s and use it in place of σ. The estimated value of σ can then be improved as more sample data are obtained, and the required sample size can be adjusted as you collect more sample data.
3. Estimate the value of σ by using the results of some other earlier study.

CHOOSING THE APPROPRIATE DISTRIBUTION



CHOOSING BETWEEN *z* AND *t*

Conditions	Method
σ not known and normally distributed population or σ not known and $n > 30$	Use Student t distribution
σ known and normally distributed population or σ known and $n > 30$	Use normal (z) distribution
Population is not normally distributed and $n \leq 30$.	Use a nonparametric method or bootstrapping

FINDING A POINT ESTIMATE AND *E* FROM A CONFIDENCE INTERVAL

Point estimate of μ:

$$\bar{x} = \frac{(\text{upper confidence limit}) + (\text{lower confidence limit})}{2}$$

Margin of error:

$$E = \frac{(\text{upper confidence limit}) - (\text{lower confidence limit})}{2}$$