\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ statistics use a process of **estimation**, estimating the value of a parameter from data obtained from a sample. For instance:

“For each dollar you pay in county property tax, 22 cents covers the cost of incarcerating prisoners.” (*Pittsburgh City Paper)*

“Eight percent of people surveyed in the United States said that they participate in skiing in the winter time” (*MRE Sports)*

These results were obtained from data collected from samples of large populations, so these values are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of the true parameters.

Common assumptions must be met before valid conclusions can be made. Some of these assumptions are:

The samples are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ selected.

The sample size must be at least \_\_\_\_\_\_\_\_\_\_ or the population must be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ . You can use a histogram to see if it is approximately \_\_\_\_\_\_\_\_\_-\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and check for \_\_\_\_\_\_\_\_\_\_\_\_\_\_.

# 7 - 1. Confidence Intervals for the Mean When σ is Known

## Objective 1. Find the Confidence Interval for the Mean When σ is Known.

### Definition: Confidence Interval

A \_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_ is a specific numerical value estimate of a parameter. The best point estimate of the population mean, σ, is the sample mean,

The sample mean is used instead of the median or mode because means of samples \_\_\_\_\_\_\_\_ less than other statistics when many samples are selected from the same population.

Good \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ have the following properties:

**Unbiased.** The expected value or the mean of the estimates obtained from samples of a given size equal to the parameter being estimated.

**Consistent.** For a consistent estimator, as sample size \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, the value of the estimator \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_the value of the parameter is estimated.

**Relatively efficient estimator.** Of all the statistics that can be used to estimate a parameter, this estimate has the smallest \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

### Confidence Intervals

Statisticians prefer to use interval estimates because the sample means are somewhat different from the population mean due to sampling error.

### Definition: Interval Estimate

An \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of a parameter is an interval or a range of values used to estimate the parameter. The estimate may or may not contain the value of the parameter being estimated.

A degree of confidence must be assigned before an interval estimate is made. This is a percentage stating how confident you want to be that the interval does contain the true population mean. The interval for 95% confidence is not as wide as the interval for 99% confidence. To increase the confidence that the true population mean is in the interval, the interval must be made \_\_\_\_\_\_\_\_\_\_\_\_\_\_.

### Definition: Confidence Level

The **confidence level** of an interval estimate of a parameter is the\_\_\_\_\_\_\_\_\_\_\_\_ that the interval estimate will contain the parameter, assuming a large number of samples are selected and that the estimation process on the same parameter is repeated.

### Definition: Confidence Interval

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is a specific interval estimate of a parameter determined by using data obtained from a sample and by using the specific confidence level of the estimate.

The interval estimate is based on the \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_ that states that when the sample size is large, approximately 95% of the sample means of same-size samples taken from a population will fall within standard errors of the population mean, that is, . (Remember that the *z-value* that separate the middle 95% from the outer 5%, 2.5% on each end, of the normal distribution are the critical values .)

When a specific sample mean, , is chosen, there is a 95% probability that the interval contains , and there is a 95% probability that the interval specified by will contain . Another way to think about this is that when intervals, , are constructed for \_\_\_\_\_\_\_ possible sample means, , 95% of the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ will contain the population mean, .

The confidence level determines the critical values of z that are used to find the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. Use Table E, or technology, to find the critical values, for different confidence intervals. α represents the total area in the two tails, so represents the area in each tail. The confidence level is . Thus, when , then

### Formula for the Confidence Interval of a Mean for a Specific α When σ is

### Known

| **Confidence Level** | **z score** |
| --- | --- |
| 90 | 1.65 |
| 95 | 1.96 |
| 99 | 2.58 |

### Definition: Margin of Error

The term is called the **margin of error** or \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. It is the likely \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ between the point estimate of a parameter and the \_\_\_\_\_\_\_\_\_\_\_\_ value of the parameter.

### The Margin of Error Depends on the Following Factors

**Population standard deviation** – the normal distribution can be used to find the confidence level when the population standard deviation is known.

**Confidence level** –if a large number of intervals are constructed, the confidence level represents the expected percentage of intervals that would contain the parameter.

**Sample size** – the margin of error decreases as the sample size increases.

### Assumptions for Finding a Confidence Interval for a Mean When α is Known.

The samples are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ selected.

The sample size must be at least \_\_\_\_\_\_\_\_\_\_ or the population must be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ .

Even with these assumptions, keep in mind there will always be a chance, α, that the confidence interval may not include the true population mean.

### Example 7.1: Number of Jobs

A sociologist found that in a random sample of 50 retired men, the average number of jobs they had during their lifetimes was 7.2. The population standard deviation is 2.1.

1. Find the best point estimate of the population mean;
2. Find the 95% confidence interval of the mean number of jobs;
3. Find the 99% confidence interval of the mean number of jobs;
4. Which is smaller? Explain why?

*Solution:*

1. The best point estimate of a population mean is the sample mean, so the best point estimate of the population mean is
2. The formula for a confidence interval is .   
    for a 95% confidence interval of the mean is 1.96, so the margin of error is   
   7.2 - .58 .58 or 6.62 7.78
3. The formula for a confidence interval is .   
    for a 99% confidence interval of the mean is 2.58, so the margin of error is

The 99% confidence interval is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

1. The \_\_\_\_\_\_% confidence interval is smaller, because

## Objective 2. Determine the Minimum Sample Size for Finding a Confidence Interval for the Mean.

### Sample Size

The answer to the question “How large a sample is necessary to make an accurate estimate?” is not simple because it depends on the margin of \_\_\_\_\_\_\_\_, the population \_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ , and the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ level.

Solve E = for *n*.

So, where E is the margin of error, then round up to the next whole number since the sample size must be a whole number.

### Example 7.2: Undergraduate GPAs

How large a sample is necessary to estimate the mean GPA of each undergraduate class at a large university within 0.25 at the 99% confidence level? The population standard deviation is 1.2.

*Solution:*

,

where E = 0.25 and σ = 1.2 and = 2.58 for the 99% confidence level.  
, when rounded up.  
A sample of 154 students is necessary to estimate the mean GPA of each undergraduate class at the university within 0.25 at 99% confidence.

Note: When using technology, use more decimal places than specified in rounding rules or shown in the examples.

# 7 – 2. Confidence Intervals for the Mean When *σ* is Unknown

## Objective 3. Find the Confidence Interval for the Mean When *σ* is Unknown.

Most of the time *σ* is unknown, so it is estimated by the sample standard deviation, *s*. When *s* is used, especially when the sample size is small, critical values greater than the values for are used in confidence intervals in order to keep the interval at any given level. These values are taken from the *Student t distribution,* most often called the ***t distribution*.**

### The *t* Distribution

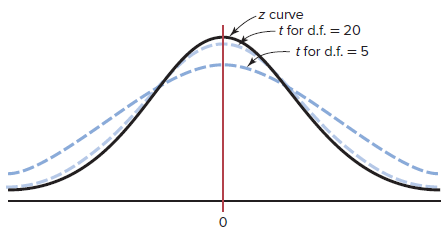
The assumptions required to use the *t distribution:*

Samples must be simple \_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

The population from which samples are selected must be normally or approximately distributed or the sample size must be \_\_\_\_\_\_ or more.

Properties of the *t* Distribution

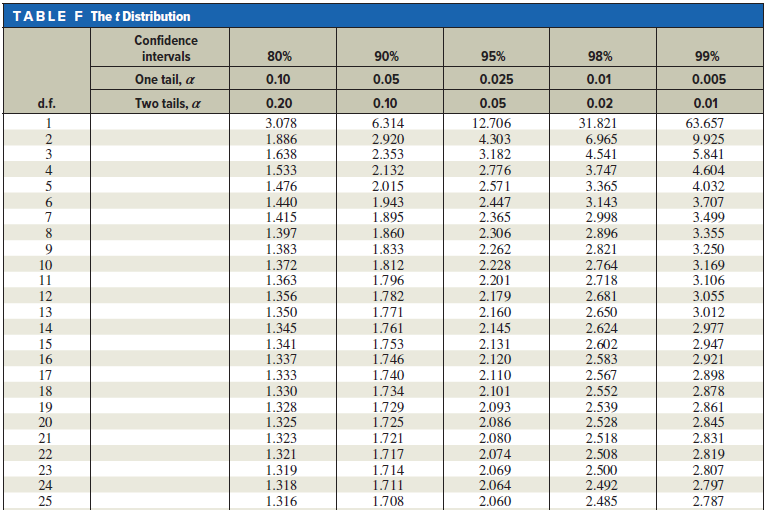
1. The *t*-distribution is \_\_\_\_\_\_-shaped.
2. The *t*-distribution is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ about the mean.
3. The mean, median, and mode are equal to \_\_\_\_\_\_\_ and are located at the \_\_\_\_\_\_\_\_\_\_\_\_ of the *t-*distribution.
4. The curve approaches, but never \_\_\_\_\_\_\_\_\_\_\_ the x-axis.
5. The standard deviation is greater than \_\_\_\_.
6. The *t* distribution is a family of curves based on \_\_\_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_\_\_\_\_, which is related to sample size.
7. As the sample size increases, the *t* distributionapproaches the \_\_\_\_\_\_\_\_\_\_\_\_\_\_ normal distribution.



Definition: Degrees of Freedom

The **degrees of freedom** are the number of values that are free to vary after a sample statistic has been computed. The degrees of freedom, d.f., for a confidence interval for the mean is found by subtracting 1 from the sample size. Thus, d.f. = \_\_\_\_\_\_\_\_\_\_\_.

Critical values, for the t distribution at specific confidence intervals and degrees of freedom are found in Appendix A, Table F, or by using technology.



Example 7-3. Finding .

Find the value for a 95% confidence interval for a sample size of 25.

*Solution:*

d. f. = 25 – 1 = 24. Find 24 in the left column labeled d.f.

Find 95% in the row labeled Confidence Intervals.

The value of

### Formula for a Specific Confidence Interval for the Mean When σ is Unknown

where the degrees of freedom are n – 1.

Example 7-4. Dance Company Students

The number of students who belong to the dance company at each of several randomly selected small universities is shown here. Estimate the true population mean size of a university dance company with 99% confidence. The data are listed below.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 21 | 25 | 32 | 22 | 28 | 30 | 29 | 30 | 47 | 26 |
| 35 | 26 | 35 | 26 | 28 | 28 | 32 | 27 | 40 |  |

*Solution:*

First, find the sample mean and standard deviation for the given sample.

\_\_\_\_\_\_\_\_\_\_\_\_, \_\_\_\_\_\_\_\_\_\_\_\_\_.

Second, find the degrees of freedom, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Next, use Table F to find the critical value, = \_\_\_\_\_\_\_\_\_\_\_\_\_\_.

And, calculate = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Finally, add the value for to and subtract it from to create the confidence interval: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Example 7-5. Hospital Noise Levels

For a random sample of 24 operating rooms taken in a hospital study, the mean noise level was 41.6 decibels, and the standard deviation was 7.5. Find the 95% confidence interval of the true mean of the noise levels in the operating rooms.

*Solution:*

# 7 – 3. Confidence Intervals and Sample Size for Proportions

## Objective 4. Find the Confidence Interval for a Proportion.

In a recent USA Today Snapshot stated that the percentage of respondents who have had someone walk into them is 36% (Source: U. S. Cellular online survey). 36% represents a proportion of cell phone users. The parameter is the proportion of cell phone users who had someone walk into them. The sample size was 738. The proportion represents the part of the whole group and can be expressed as a fraction, decimal, or percentage or represent a probability. For instance, if a cell phone user is randomly selected, the probability that they have had someone walk into them is 0.36.

### Symbols Used in Proportion Notation

For a sample proportion. , , or ,  
where *X* = number of sample units that possess the characteristics of interest and *n* = sample size.

### Example 7 – 6. Find the Sample Proportion and its Complement

For *n* = 80 and *X* = 45, find and .

*Solution:*

or .

### Confidence Intervals

### The assumptions that must be made to find a confidence interval for a population proportion are:

The sample must be a random sample.

The conditions for a binomial experiment must be met.

The \_\_\_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_ for proportions is and and .

### Formula for a Specific Confidence Interval for a Proportion

### Example 7 – 7. Travel to Outer Space

A CBS News/*New York Times* poll found that 329 out of 763 randomly selected adults said they would travel to outer space in their lifetime, given the chance. Estimate the true proportion of adults who would like to travel to outer space with 92% confidence.

*Solution:*

and

Look up the value of for, , that is, in Appendix A, Table E.

### Example 7 – 8. Cell Phone Poll

In a poll of 738 cell phone users, 266 stated that they had been walked by someone using their cell phone. Estimate the true proportion of cell phone users who have had someone walk into them with 95% confidence.

*Solution:*

## Objective 5. Determine the Minimum Sample Size for Fining a Confidence Interval for a Proportion.

### Sample Size

To find the sample size needed to determine a confidence interval about a proportion, we solve the margin of error formula for the sample size, n:

So, . Remember to round up to the nearest whole number, if necessary.

### Example 7 – 9. Widows

A recent study indicated that 29% of the 100 women over age 55 in the study were widows. How large a sample is needed to be 95% confident that the estimate is within .03 of the true proportion of women over age 55 who are widows?

*Solution:*

29% of 100 = 29 = *X*

*E* = 0.03

For , .

rounded up is \_\_\_\_\_\_\_\_\_.

The minimum sample size needed to estimate the true proportion of widows within 3% at a 95% confidence level is \_\_\_\_\_\_\_\_.

### Example 7 – 10. Diet Habits

A federal report indicated that 27% of children ages 2 to 5 years had a good diet. How large a sample is needed to estimate the true proportion of children 2 to 5 with good diets with 2% with 98% confidence?

*Solution:*

The minimum sample size needed to estimate the true proportion of children 2 to 5 with good diets within 2% at a 98% confidence level is \_\_\_\_\_\_\_\_.

# 7 – 4. Confidence Intervals for Variances and Standard Deviations

## Objective 6.Find a Confidence Interval for a Variance and a Standard Deviation.

In manufacturing, it is important that products dimensions are consistent so that they will fit together or that the amount of active ingredient in a medication have little variation so that it delivers the correct dosage. Thus, confidence intervals for variances and standard deviations are necessary.

A new distribution, called the \_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_ distribution, is used.

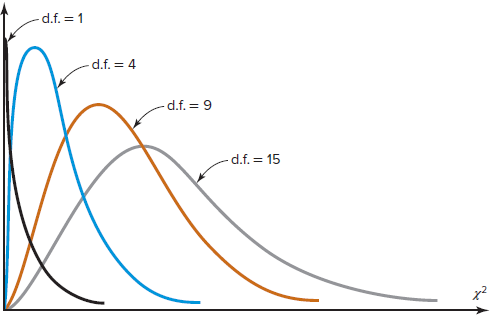
### Characteristics of the Chi-Square Distribution

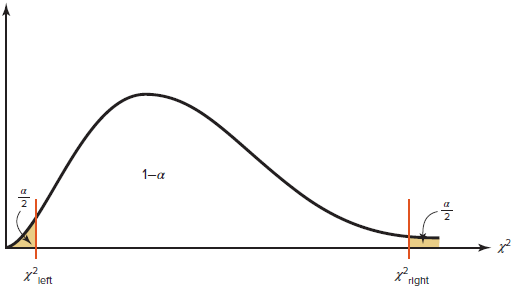
1. All chi-square values are greater than or equal to \_\_\_\_\_\_\_\_\_\_.

2. The chi-square distribution is a family of curves based on the \_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_\_\_\_\_.

3. The area under each chi-square distribution is equal to \_\_\_\_\_\_\_.

4. The chi-square distributions are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ skewed.

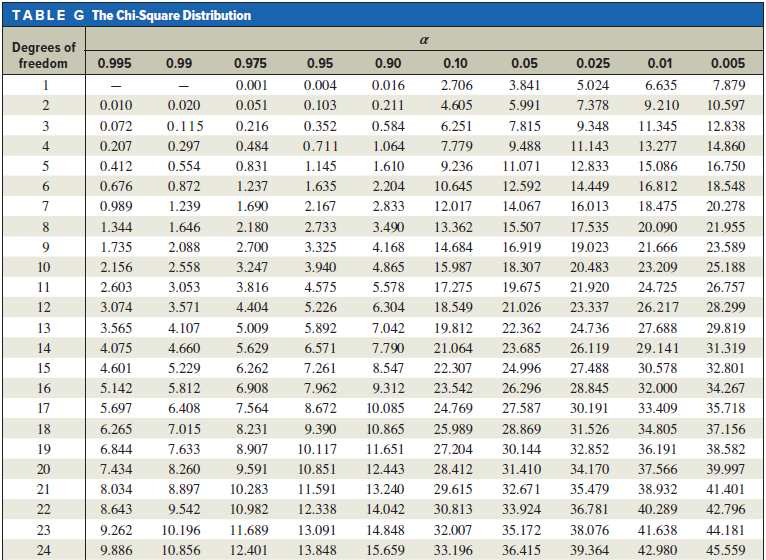




To find the critical values of χ2 use Table G in Appendix A, you need to know the \_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, (), the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ level, and the value of . For instance, for a confidence level of 98, .

Locate the degrees of freedom in the left hand column and the values for and in the top row. The value of is in the row corresponding to the degrees of freedom and column corresponding to . The value of is in the row corresponding to the degrees of freedom and column corresponding to .

If the number of degrees of freedom is not given in the table, use the closest \_\_\_\_\_\_\_\_\_\_ value in the table.



### Example 7.11. Find and

Find and for and .

*Solution:*

Degrees of freedom =

From Table G, .

From Table G,

Find and for and .

*Solution:*

Degrees of freedom = \_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

From Table G, \_\_\_\_\_\_\_\_\_\_\_

From Table G, \_\_\_\_\_\_\_\_\_\_

### Formula for the Confidence Interval for a Variance

where

### Formula for the Confidence Interval for a Standard Deviation

\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_ where

### Assumptions for Finding a Confidence Interval for a Variance or Standard Deviation

The sample is a \_\_\_\_\_\_\_\_\_\_\_\_ sample.

The population must be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

When computing a confidence interval for a population variance or \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ using raw data, round off to one more decimal place that the number of decimal places in the original data. However, if you are using sample statistics, round off to the \_\_\_\_\_\_\_\_\_\_\_\_\_ number of decimal places as given for the sample variance or sample standard deviation.

### Example 7.12. Lengths of Children’s Animated Films

The lengths (in minutes) of a random selection of popular children’s animated films are listed below. Assume the variable is normally distributed. Estimate the population variance and standard deviation in length of children’s animated films with 99% confidence.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 93 | 83 | 76 | 92 | 77 | 81 | 78 | 100 | 78 | 76 | 75 |

*Solution:*

Find the variance using technology. \_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_, ­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_, and \_\_\_\_\_\_\_\_\_\_

Find \_\_\_\_\_\_\_\_\_and \_\_\_\_\_\_\_\_\_\_\_\_\_

Substitute into the formula:

Now, take the square root of each side to find the confidence interval about the standard deviation:

### Example 7.13. Lifetimes of Wristwatches

Find the 90% confidence interval for the variance and standard deviation for the lifetimes of inexpensive wristwatches if a random sample of 24 watches has a standard deviation of 4.8 months. Assume the variable is normally distributed. Do you believe that the lifetimes are relatively consistent? Why or why not?

*Solution:*

\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_, ­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_, and \_\_\_\_\_\_\_\_\_\_

Find \_\_\_\_\_\_\_\_\_and \_\_\_\_\_\_\_\_\_\_\_\_\_

Substitute into the formula:

(Yes or No), the lifetimes (are or are not) relatively consistent, because