Researchers are interested in answering many types of questions. These questions are stated as claims about a population and the decision making process for evaluating these claims is called \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. Researchers define the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ under study, state the hypothesis or claim to be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, set the significance level, \_\_\_\_\_\_\_\_\_\_\_\_ a sample from the population, \_\_\_\_\_\_\_\_\_\_ the data, run the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ test, and reach a conclusion.

We will consider three methods of performing \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_:

1. The traditional method;
2. The *P*-value method; and
3. The confidence interval method.

# 8 - 1. Steps I Hypothesis Testing – Traditional Method

## Objective 1. Understand the Definitions Used in Hypothesis Testing.

### Definition: Statistical Hypothesis

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is a conjecture abut a population parameter. The conjecture may or may not be \_\_\_\_\_\_\_\_\_\_\_\_.

There are always two sides to a conjecture – the conjecture and its complement.

### Definition: Null Hypothesis

The **null hypothesis**, symbolized by \_\_\_\_\_\_, is a statistical hypothesis that states that there is no difference between a parameter and a specific value, or that there is no \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ between two parameters.

### Definition: Alternate Hypothesis

The **alternative hypothesis**, symbolized by \_\_\_\_\_\_\_\_\_, is a statistical hypothesis that states the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of a difference between a parameter and a specific value or states that there is a difference between two parameters.

The null hypothesis always states that the parameter is equal to a specific value or to another parameter. The alternative indicates change, so the alternative hypothesis claim that the parameter is either not equal to, greater than, or less than a specific value or than another parameter.

## Objective 2. State the Null and Alternate Hypotheses.

### Example 8 - 1. State the Null and Alternative Hypotheses

1. The average age of first-year medical school students is less than 29 years.
2. The average number of seasons in which an NBA player participates is 4.71.
3. The average number of monthly sessions on the Internet by a person at home has increased from 36.
4. The average cost of a basic cell phone plan is at least $79.95.
5. The average weight loss for a sample for a sample of people who exercise 30 minutes per day for 6 weeks is 8.2 pounds.

*Solution:*

1. The average age of first-year medical school students has decreased from 29 years.

The researcher knows a previous value for the average age and believes that the age has decreased. Thus, no change would mean that the age is the same as it was before. The\_\_\_\_\_\_\_\_\_ hypothesis is that the average age is 29 years, *H*0: . The \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis is that the average age is less than 29 years, *H*1: .   
This test indicates that the only change of interest is less than 29, so this test is called *a left-tailed test.*

1. The average number of seasons in which an NBA player participates is 4.71 seasons.

The researcher is concerned with the number of seasons NBA players play. Has the number of seasons increased, decreased, or remained unchanged? A previous study indicates the average number of seasons for the population of NBA players. The hypotheses for this claim are   
*H*0: and *H*1: .

This test indicates that change could be more than or less than 4.71 seasons and is called a *two-tailed test.*

1. The average number of monthly sessions on the Internet by a person at home has increased from 36.

*H*0: and *H*1: .   
The change of interest is an increase from 36 or a number of sessions greater than 36. This test is called a *right-tailed test.*

1. The average cost of a basic cell phone plan is at least $79.95.

The words “at least” mean “equal to or greater than.” Thus, change would mean that the average cost is less than $79.95. *H*0: and *H*1: . This is a \_\_\_\_\_\_\_\_\_\_\_-tailed test.

1. The average weight loss for a sample of people who exercise 30 minutes per day for 6 weeks is 8.2 pounds.   
   *H*0: and *H*1: . This is a \_\_\_\_\_\_\_\_\_\_\_-tailed test.

| **Two-tailed Test** | **Right-tailed test** | **Left-tailed test** |
| --- | --- | --- |
| *H*0: | *H*0: | *H*0: |
| *H*1: | *H*1: | *H*1: |

The null hypothesis is always stated using the \_\_\_\_\_\_\_\_\_\_\_\_\_ sign!

When researchers conduct a study, they are usually looking for evidence to support their claim. Thus, the claim should be stated as the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis. The alternative hypothesis is also called the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis.

A claim can be either the \_\_\_\_\_\_\_\_\_ or the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis.

However, statistical evidence only *supports* the claim if it is the alternative hypothesis. Statistical evidence can be used to *reject* the claim if the claim is the null hypothesis.

After stating the hypothesis, the study is designed. The researcher selects the correct *statistical test*, chooses an appropriate *significance level*, and makes a plan for conducting the study.

In example 8-1, part e, the researcher selects a sample of patients who will exercise 30 minutes a day for 6 weeks. At the beginning, and again at the end, of the 6 weeks, the weights of the people are measured.

Recall that for samples of a specific size randomly selected from a population, the means of the samples will \_\_\_\_\_\_\_\_ about the population mean, and the distribution of the sample means, for sample sizes of 30 or more, will be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. Even if the weight loss is exactly 8.2 pounds and the null hypothesis is true, the mean of the weight loss for a sample of people will not, in most cases be equal to the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ of 8.2 pounds.

There are two possibilities:

The null hypothesis is *true* and the difference between the population mean and the sample mean is due to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Or

The null hypothesis is *false* and the sample came from a population came from a population whose mean weight loss is not 8.2 pounds, but is some other value that is unknown.

The \_\_\_\_\_\_\_\_\_\_\_\_\_ the difference between the claimed population mean and the sample mean, the more evidence there is for rejecting the null hypothesis. The probability that the sample came from a population with a mean of 8.2 decreases as the distance between the means \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

The decision must be made \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, rather than based on feelings or intuition. The difference must be significant and most likely not be due to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

### Definition: Statistical Test

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_ uses data obtained from a sample to make a decision about whether the null hypothesis should be rejected.

### Definition: Test Statistic

The test value or \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_ is the numerical value obtained from a statistical test.

Decisions made to reject or not to reject the null hypothesis on the basis of the test statistic obtained from a \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_ rely on data from a sample drawn from the population. If the difference between the sample mean and the population mean is deemed to be significant, the hypothesis is rejected. Otherwise, the null hypothesis is not rejected.

Because sample data are used to decide to reject a null hypothesis or not, it is possible that an incorrect decision can be made.

In reality, the null hypothesis may actually be \_\_\_\_\_\_ or may not actually be \_\_\_\_\_\_\_\_. Regardless of actual truth, a decision to reject the null hypothesis or not is made. Thus, there are two possibilities for a \_\_\_\_\_\_\_\_\_\_\_\_ decision and two for an \_\_\_\_\_\_\_\_\_\_\_\_\_ decision:

1. We reject the null hypothesis when it is true. This is an incorrect decision and results in a **Type I error**.
2. We reject the null hypothesis when it is false. This is a correct decision.
3. We do not reject the null hypothesis when it is true. This is a correct decision.
4. We do not reject the null hypothesis when it is false. This is an incorrect decision and results in a **Type II error**.

|  | H0 is true | H0 is false |
| --- | --- | --- |
| Reject H0 | Incorrect  **Type I Error** | Correct decision |
| Do not reject H0 | Correct decision | Incorrect  **Type II Error** |

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ occurs if a true null hypothesis is rejected.

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ occurs if a false null hypothesis is not rejected.

### Example 8 – 2. Medical Research

A medical researcher is interested in finding out if a new medication changes the pulse rate of patients who take it. Suppose the mean pulse rate for the population understudy is 82 beats per minute. Then, the null hypothesis is that there is no change in the patients’ mean pulse rate from 82 beats per minute after taking the medication: . The alternate hypothesis is that the mean pulse rate for patients after taking the medication is not 82 beats per minute: . Identify the decisions that would be correct and which ones would be Type I or Type II errors:

*Solution:*

If the null hypothesis is, in reality, true, then it is a correct decision to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis. In other words, the mean pulse rate for the population of patients is 82 beats per minute, and the sample mean is not significantly different from 82 beats per minute. The researcher believes that the medication does not change patients’ pulse rates. The researcher does not reject a true null and made a \_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

If the null hypothesis is, in reality, true, then it is an incorrect decision to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis. This would be a Type I error. In other words, the mean pulse rate for the population of patients is 82 beats per minute, but the sample mean is significantly different from 82 beats per minute. The researcher believes that the medication changes patients’ pulse rates. The researcher has rejected a true null hypothesis and made a \_\_\_\_\_\_\_\_\_\_\_\_\_\_ error.

If the null hypothesis is, in reality, false, then it is a correct decision to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis. In other words, the mean pulse rate for the population of patients is not really 82 beats per minute, but the sample mean is significantly different from 82 beats per minute. The researcher believes that the medication changes patients’ pulse rates. The researcher has rejected a false null and made a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

If the null hypothesis is, in reality, false, then it is an incorrect decision to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis. This is a Type II error. In other words, the mean pulse rate for the population of patients is actually not 82 beats per minute, but the sample mean is not significantly different from 82 beats per minute. The researcher believes that the medication does not change patients’ pulse rates. The researcher has not rejected a false null and made a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ error.

Researchers make decisions based on sample data. If the evidence against the null hypothesis is strong enough, the null hypothesis is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. If the evidence is weak, the null hypothesis is not rejected. Their decisions do not prove that the null hypothesis is true or false. Statistics do not \_\_\_\_\_\_\_\_\_\_\_ anything absolutely. The only way to prove anything statistically is to use the entire \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Thus, it is important to decide how large the difference between the mean obtained from the sample and the hypothesized mean must be in order to decide whether to reject the null hypothesis or to not reject the null hypothesis.

### Definition: Significance Level.

The \_\_\_\_\_\_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, symbolized by α, is the maximum acceptable probability of making a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.   
That is, .

Statisticians usually agree on using three arbitrary significance levels, 0.10, 0.05 and 0.01. Depending on the level of significance selected, if the null hypothesis is rejected, the probability of making a type 1 error is \_\_\_\_\_\_\_\_%, \_\_\_\_\_\_\_\_\_%, or \_\_\_\_\_\_\_\_\_\_%.

## Objective 3. Find the Critical Values for the z-Test.

Once a significance level is chosen, the corresponding critical value is selected from a table for appropriate test. If a z test is used, the *z*-table (Table E) is used. The critical value is the value of *z* that separates critical region(s) from non-critical region(s).

### Definition: Critical Region

The **critical** or **rejection region** is the \_\_\_\_\_\_\_\_\_\_\_ of test values that indicates that there is a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ difference and that the null hypothesis should be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

### Definition: Noncritical Region

The **noncritical** or **nonrejection region** is the range of test values that indicates that the difference was probably due to \_\_\_\_\_\_\_\_\_\_\_\_ and that the null hypothesis should \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

### Definition: Critical Value

The **critical value** separates the critical region from the noncritical region. The symbol for critical value is C.V.

For an alternative hypothesis *H*1: μ > 40 that uses the \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ symbol, the null hypothesis will be rejected only when the sample has a mean that is significantly greater than 40. Thus the critical value must be on the right side of the mean. Then the test is called a \_\_\_\_\_\_\_\_\_\_-\_\_\_\_\_\_\_\_\_\_\_\_ test.

For an alternative hypothesis *H*1: μ < 40 that uses the \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_ symbol, the null hypothesis will be rejected only when the sample has a mean that is significantly less than 40. Thus the critical value must be on the left side of the mean. Then the test is called a \_\_\_\_\_\_\_\_\_\_-\_\_\_\_\_\_\_\_\_\_\_\_ test.

### Definition: One-Tailed Test

A \_\_\_\_\_\_\_-\_\_\_\_\_\_\_\_\_ test indicates that the null hypothesis should be rejected when the test value is in the \_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_ on one side of the mean. A one-tailed test is either a \_\_\_\_\_\_\_\_\_\_-tailed test or a \_\_\_\_\_\_\_\_\_\_-tailed test, depending on the direction of the inequality of the alternative hypothesis.

To find the critical value, the researcher must choose a significance level, called an alpha (*α*) level. Remember that *α* is the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of the confidence level.

Suppose the medical researcher chose to use a 99% confidence level, and thus, an alpha of . If the alternate hypothesis is that the mean pulse rate is greater than 82 beats per minute, *μ* > 82, then the *z* value is such that 1% of the area falls to the \_\_\_\_\_\_\_ and 99% of the area falls to the \_\_\_\_\_\_\_\_\_\_\_\_\_. Using Table, E, find the area value in the body of Table E closest to 99% or 0.9900 (in this case, 0.9901). Thus the value of *z* = 2.33.

The normal curve is shown with the mean at z = 0 and the critical value of z to the right.  The area under the normal curve to the right of the critical value is shaded and labeled as 0.01.  This is the critical region.   The table is shown for the area  closest to 0.9900, which is 0.9901.  The z value associated with the area of 0.9901 is 2.33 because 0.9901 is in the row for a z value of 2.3 and column for the second deciaml place of 0.03.

Suppose the medical researcher chose an alternative hypothesis that the mean pulse rate is less than 82 beats per minute, *μ* < 82, then the *z* value is such that 1% of the area falls to the \_\_\_\_\_\_\_\_\_\_\_\_ and 99% of the area falls to the \_\_\_\_\_\_\_\_\_\_\_\_ This is a left-tailed alternative hypothesis. Using Table, E, find the area value closest to 1% or 0.0100 (in this case, 0.0099). The value of *z* is −2.33.

When the research question is about whether the mean is different from a particular value, as in the case of an alternative hypothesis that the mean pulse rate is not equal to 82 beats per minute, *μ* ≠ 82. The researcher is conducting a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ test and the null hypothesis can be rejected when there is a significant \_\_\_\_\_\_\_\_\_\_\_\_\_ above or below the mean.

For a two-tailed test, the critical region is separated into two equal parts, each having an area of one-half of alpha. For instance, when *α* = 0.01, then and 0.005 is to the right of the mean and 0.005 is to the left of the mean. The *z* score to the left is found by looking up the *z* score corresponding to the area of 0.005 on Table E. 0.005 falls halfway between 0.0051 and 0.0049 on the table. The *z* score is if the *z* value is needed to three decimal places, or −2.58 if rounded to two decimal places. The value of the *z* score on the right is the opposite of the left-hand *z* score, due to \_\_\_\_\_\_\_\_\_\_\_ and, thus, is either 2.575 or 2.58.

The normal curve is drawn and the mean of 0 is marked in the center.  The left and right sides are identical.  Then the critical region of 0.005 (half of 0.01) is marked on the left and the right.  The z scores of -2.58 and +2.58 (rounded to two decimal places) label the intersection of the marks for the critical regions and the horizontal axis.

### Steps for finding Critical Values for Specific α Values, Using Table E.

**Step 1**

Draw the figure (the normal curve) and indicate the appropriate area.

1. If the test is left-tailed, the critical region, with an area equal to *α*, will be on the \_\_\_\_\_\_\_\_ side of the mean.
2. If the test is right-tailed, the critical region, with an area equal to *α*, will be on the \_\_\_\_\_\_\_\_\_\_ side of the mean.
3. If the test is two-tailed, *α* must be divided by 2, one-half of the area will be to the \_\_\_\_\_\_ of the mean and one-half of the area will be to the \_\_\_\_\_\_ of the mean.

**Step 2**

1. For a left-tailed test, use the *z* value that corresponds to the area equivalent to α in Table E.
2. For a right-tailed test, use the *z* value that corresponds to the area equivalent to (the \_\_\_\_\_\_\_\_\_\_of *α*) in Table E.
3. For a two-tailed test, use the *z* value that corresponds to for the left value. It will be negative. For the\_\_\_\_\_\_\_\_ value, use the *z* value that corresponds to the area equivalent to (or use the opposite of the left value, due to symmetry). It will be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

## Objective 4. State the Five Steps Used in Hypothesis Testing.

### Steps for Hypothesis Testing

1. State the null and alternative hypotheses and identify the claim.
2. Design the study. Select the correct statistical\_\_\_\_\_\_, choose a level of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, define the population, determine how the \_\_\_\_\_\_\_\_\_\_\_ is selected, and decide the methods used to \_\_\_\_\_\_\_\_\_\_\_\_ the data.
3. Conduct the study and \_\_\_\_\_\_\_\_\_\_\_\_\_\_ the data.
4. Evaluate the data. Tabulate the data, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the statistical test, decide whether to reject or not reject the \_\_\_\_\_\_\_ hypothesis.
5. Summarize the\_\_\_\_\_\_\_\_\_\_\_\_\_.

### Solving Hypothesis Testing Problems (Using the Traditional Method)

1. \_\_\_\_\_\_\_\_\_\_\_\_\_ the hypotheses and identify the claim.
2. Find the \_\_\_\_\_\_\_\_\_\_\_\_\_value(s) from the appropriate table.
3. Compute the \_\_\_\_\_\_\_\_\_ value.
4. Make the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ to reject or not reject the null hypothesis.
5. \_\_\_\_\_\_\_\_\_\_\_\_\_\_ the results.

### Summarize the Results of a Hypothesis Test

|  | **Claim** | |
| --- | --- | --- |
| **Decision** | **Claim is Null Hypothesis (*H*0)** | **Claim is Alternative Hypothesis (*H*1)** |
| Reject *H*0 |  |  |
| Do not Reject *H*0 |  |  |

### Statistical Decisions Are Not Proof

Statisticians only state that there is or is not enough evidence to conclude that a claim is *probably* \_\_\_\_\_\_\_ or \_\_\_\_\_\_\_\_\_\_\_. Nothing is actually \_\_\_\_\_\_\_\_\_. The only way to prove something would be to use the \_\_\_\_\_\_\_\_\_\_ population. It is usually \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ or impractical to do this, especially when the population is large.

### Example 8 – 3. Eggs and Your Health

The Incredible Edible Egg Company recently found that eating eggs does not increase a person’s blood serum cholesterol. Five hundred American adults participated in a study that lasted for 2 years. The participants were randomly assigned to either a no-egg group or a moderate-egg group. The blood serum cholesterol levels were checked at the beginning and at the end of the study. Overall the groups’ levels were not significantly different. The company reminds us that eating eggs is healthy if done in moderation. Many of the previous studies relating eggs and high blood serum cholesterol jumped to improper conclusions.

1. Why did the company conduct the study?
2. Identify the population under study.
3. What is the research hypothesis? Is it the null or alternative hypothesis?
4. Was a sample collected?
5. Were data collected? If so, what data?
6. Were statistical tests run? If so, were they left-, right-, or two-tailed tests?
7. What conclusion was drawn?
8. What might be the reason that previous studies drew different conclusions?

*Solution:*

1. Why did the company conduct the study?
2. Identify the population under study.
3. What is the research hypothesis? Is it the null or alternative hypothesis?
4. Was a sample collected?
5. Were data collected? If so, what data?
6. Were statistical tests run? If so, were they left-, right-, or two-tailed tests?
7. What conclusion was drawn?
8. What might be the reason that previous studies drew different conclusions?

# 8 – 2. *z* Test for a Mean

## Objective 5. Test Means When *σ* is Known, Using the *z* Test.

The statistical test called the *z* test for a mean is used when *σ* (the population standard deviation) is known. When we do not know the population standard deviation, we cannot use the *z* test and must use a different test which is discussed in the next section.

### The formula for the test statistic is

where the **observed value** is the \_\_\_\_\_\_\_\_\_\_\_\_, such as the sample mean, computed from the sample data; the **expected value** is the \_\_\_\_\_\_\_\_\_\_\_\_\_\_, such as the population mean that is expected if the null hypothesis is \_\_\_\_\_\_\_; and the denominator is the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, such as the standard error of the mean.

Thus the *z* test is a statistical test for the mean of a population. It is appropriate when the sample is a random sample, either or the population is \_\_\_\_\_\_\_\_\_\_\_\_ distributed and *σ* is known.

where

The first step for testing hypotheses is to check that the assumptions have been met. During this class the assumptions are stated in the exercises.

### Recall the Five Steps for Solving Hypothesis Testing Problems:

**Step 1 –** State the null and alternative hypotheses and identify the claim.

**Step 2 –** Find the critical value(s).

**Step 3 –** Compute the \_\_\_\_\_\_\_\_\_ statistic.

**Step 4 –** Make the decision: \_\_\_\_\_\_\_\_\_ or \_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.

**Step 5 –** Summarize the results.

A sample statistic may be lower (or higher) than a hypothesized parameter without being significantly lower or higher. In other words, the difference may be due to \_\_\_\_\_\_\_\_\_\_\_\_. However, there is still a chance of a Type II error, or not rejecting a null hypothesis when it is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

When the difference is statistically significant, the null hypothesis is rejected. Significance is the \_\_\_\_\_\_\_\_\_\_\_\_\_\_ of making a Type I error the researcher has set, so keep in mind that when a null hypothesis is rejected, there is a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of making a Type I error and rejecting a null hypothesis that is \_\_\_\_\_\_\_\_\_\_\_\_.

When a null hypothesis is not rejected, you may not accept it as true. There is not enough evidence to say that it is false.

When a null hypothesis is rejected, you may not say it is false. There is, however, enough evidence to conclude that it is false.

## Using the Traditional Method for Hypothesis Testing

### Example 8 – 4. *z* test Using the Critical Value Method

The average 1-year-old (both genders) is 29 inches tall. A random sample of 30 1-year-olds in a large day care franchise resulted in the following sample of heights. Can it be concluded that the average height differs from 29 inches at a significance level of 0.05? Assume *σ* = 2.61 inches.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 25 | 32 | 35 | 25 | 30 | 26.5 | 26 | 25.5 | 29.5 | 32 |
| 30 | 28.5 | 30 | 32 | 28 | 31.5 | 29 | 30 | 34 | 29.5 |
| 29 | 32 | 27 | 28 | 27 | 32 | 29.5 | 33 | 28 | 29 |

*Solution:*

The null hypothesis states \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is 29 inches.

The alternate hypothesis states \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 29 inches.

*H*0: *µ* \_\_\_\_\_\_\_\_ *H*1: *µ* \_\_\_\_\_\_\_\_

The claim is the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis.

The significance level is 0.05.

The test is \_\_\_\_\_\_-tailed. The critical value(s) is/are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Using technology, find the test statistic. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Compare the critical value(s) and direction to the test statistic in order to make a decision (Choose the appropriate decision.):

Reject the null hypothesis. Do not reject the null hypothesis.

There (is / is not) enough evidence to support the claim that the average number of 1-year olds \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**Remember**: When the test statistic falls in the critical region, your decision is to reject the null hypothesis. It indicates that the difference is \_\_\_\_\_\_ likely to have been due to \_\_\_\_\_\_\_\_\_\_.

**Remember:** When the test statistic falls in the non-critical region, the decisions is to\_\_\_\_\_\_\_ reject the null hypothesis. It indicates that the difference is likely due to \_\_\_\_\_\_\_\_\_\_.

## Using the *P*-Value Method for Hypothesis Testing

Usual levels of significance used to test hypotheses are 0.05, \_\_\_\_\_\_\_, and sometimes \_\_\_\_\_\_\_\_\_. The choice depends of how \_\_\_\_\_\_\_\_\_\_\_\_\_\_ it is that we do not make a Type I error.

Use of technology has made finding another statistic, the *P*-value for the hypothesis test.

### Definition: *P*-value

The ***P*-value** (probability level) is the probability of getting a sample statistic (such as the sample mean) or a more extreme sample statistic in the direction of the alternative hypothesis when the null hypothesis is true.

In other words, the *P*-value is the actual \_\_\_\_\_\_\_\_\_ under the standard normal distribution curve (or other curve, depending on what statistical test is used) representing the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of a particular sample statistic or a more extreme sample statistic occurring if the null hypothesis is \_\_\_\_\_\_\_\_\_\_.

### Example 8 – 5. Decision for a *P*-value for Different Significance Levels

Suppose a computer printout shows a *P*-value of 0.0356.

For an alternative hypothesis *H*1: µ > 50, at *α*the null hypothesis will be rejected since the *P*-value is less than *α*. However, if the same hypothesis test used *α* = 0.01, the null hypothesis is not rejected because the *P*-value is greater than *α*.

*Solution:*

If the hypothesis test is two-tailed, the area in one tail must be doubled. For a two-tailed test with *α* = 0.05 and *P*-value = 0.0356, the *P*-value for two tails will be 2(0.0356) = 0.0712. Thus, the null hypothesis will not be rejected because the *P*-value is greater than *α*.

### Decision Rule When Using a P-Value

If *P*-value , reject the null hypothesis.

If *P*-value , do not reject the null hypothesis.

Notice that *α* is chosen by the researcher *before* the statistical test is conducted while the *P*-value is computed *after* the sample mean (or test statistic) has been found.

For a *z* test, you can use Table E in Appendix A to find the *P*-values (or use technology). First, find the area under the standard normal distribution curve corresponding to the *z* test value. For a left-tailed test, use the area given in the table. For a right-tailed test, use 1.0000 minus the area given in the table. To find the *P*-value for a two-tailed test, double the area you found in one tail. However, calculators and computers software that calculate the *P*-value calculate the actual *P*-Value, so subtraction and doubling are not needed.

### Procedure Table. Solving Hypothesis Testing Problems using the P-Value Method

**Step 1 –** State the null and alternative hypotheses and identify the claim.

**Step 2 –** Compute the \_\_\_\_\_\_\_\_\_ statistic.

**Step 3 –** Find the *P*-Value.

**Step 4 –** Make the decision: \_\_\_\_\_\_\_\_\_ or \_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.

**Step 5 –** Summarize the results.

### Guidelines for *P*-Values

* 1. If *P*-value reject the null hypothesis. The difference is highly significant.
  2. If *P*-value but *P*-value reject the null hypothesis. The difference is significant.
  3. If *P*-value but *P*-value consider the consequences of Type I error before rejecting the null hypothesis.
  4. If *P*-value do not reject the null hypothesis. The difference is not significant.

### Example 8 – 6. Soft Drink Consumption

A researcher claims that the yearly consumption of soft drinks per person is more than 52 gallons. In a sample of 50 randomly selected people, the mean of the yearly consumption was 53.7 gallons. The population standard deviation is 3.5 gallons. Find the *P*-value for the test. (You may use technology or the table.) On the basis of the *P*-value, is the researcher’s claim valid?

*Solution:*

1. The null hypothesis is that \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
The alternate hypothesis, the researcher’s claim is that \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ This is a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_-tailed test.

1. The test statistic for the *z*-test is .
2. The *P*-Value is \_\_\_\_\_\_\_\_\_\_\_\_\_\_ from the Table E. The *P*-Value found using technology is .
3. The decision is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.
4. There is sufficient evidence to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. The test \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the decision that the researcher’s claim is valid.

# 8 – 3. *t* Test for a Mean

## Objective 6. Test Means When *σ* is Unknown, Using the *t* Test.

More often than not, the standard deviation of the population is unknown, so the *z* test is not typically used for testing hypotheses involving means. The *t* test is used instead. The distribution of the variable should be approximately normal.

### Properties of the *t*-distribution similar to the standard normal 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 distribution.
  4. The curve approaches but never \_\_\_\_\_\_\_\_\_\_\_ the *x*-axis.

### Properties of the *t*-distribution that differ from the standard normal distribution

1. The standard deviation is greater than \_\_\_\_.
2. The *t distribution* is a family of curves based on \_\_\_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_\_\_\_\_, which is related to sample size.
3. As the sample size increases, the *t distribution* approaches the \_\_\_\_\_\_\_\_\_\_\_\_\_\_ normal distribution.

### Definition: The *t* test

The *t* test is a statistical test for the mean of a population and is used when the population is normally or approximately normally distributed and the population standard deviation, *σ*, is unknown.

The formula for the *t* test is

The degrees of freedom are d.f. = *n* – 1.

Use Table F to find the critical value(s) for the *t* test. Remember, for degrees of freedom larger than 30, round down to the nearest table value for the conservative approach.

As degrees of freedom get larger, the critical values for the *t* test approach the *z* values.

### Assumptions for the *t* test for a Mean When σ is Unknown

1. The sample is a random sample.
2. Either or the population is normally distributed when

### The Procedure for the *t* Test by the Traditional Method:

**Step 1 –** State the null and alternative hypotheses and identify the claim.

**Step 2 –** Find the critical value(s) using Table F.

**Step 3 –** Compute the \_\_\_\_\_\_\_\_\_ statistic.

**Step 4 –** Make the decision: \_\_\_\_\_\_\_\_\_ or \_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.

**Step 5 –** Summarize the results.

### The Procedure for the t Test by the *P*-value Method:

**Step 1 –** State the null and alternative hypotheses and identify the claim.

**Step 2 –** Compute the \_\_\_\_\_\_\_\_\_ statistic.

**Step 3 –** Find the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**Step 4 –** Make the decision: \_\_\_\_\_\_\_\_\_ or \_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.

**Step 5 –** Summarize the results.

### Example 8 – 7. Use Table F to Find the P-value Interval for Each Test Value

1. *t* = 2.321, n = 15 right-tailed
2. *t* = −1.415, n = 24, two-tailed
3. *t* = −1.862, n = 17, left-tailed

t table for degrees of freedom from 1 through 65 and significance levels for one and two tailed tests with confidence intervals of 80%, 90%, 95%, 98%, and 99%.

*Solution:*

1. In Table F, choose the row for degrees of freedom equal to . 2.321 = *t* is between table values of \_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_. The one-tailed areas above these values are \_\_\_\_\_\_\_ and \_\_\_\_\_\_\_. Therefore, \_\_\_\_\_\_ < *P*-value < \_\_\_\_\_. (The *P*-value obtained using a computer or calculator is \_\_\_\_\_\_\_\_\_\_.)
2. In Table F, choose the row for degrees of freedom equal to . Notice that 1.415 is between \_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_. The table shows only positive values, but we can consider that −1.415 is between \_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_ as well. For a two-tailed test, the *P*-value is between \_\_\_\_\_ and 0\_\_\_\_\_\_. That is, . (The *P*-value obtained using a computer or calculator is \_\_\_\_\_\_\_\_.)
3. In Table F, choose the row for degrees of freedom . In row 16, we find the value of *t* = −1.862 is between \_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_. Thus, for a left-tailed test, . (The *P*-value obtained using a computer or calculator is \_\_\_\_\_\_\_\_.)

### Example 8−8. Chocolate Chip Cookie Calories

The average 1-ounce chocolate chip cookie contains 110 calories. A random sample of 15 different brands of 1-ounce chocolate chip cookies resulted in the following calorie amounts. At the α = 0.01 level, is there sufficient evidence that the average calorie content is greater than 110 calories?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 100 | 125 | 150 | 160 | 185 | 125 | 155 | 145 |
| 100 | 150 | 140 | 135 | 120 | 110 | 160 |  |

*Solution:*

**Step 1.**

**Find the sample mean and sample standard deviation:**

In order to calculate the test statistic, we must find the sample mean and standard deviation for the sample.

**Step 2:**

The test statistic is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Step 3:**

The critical value is \_\_\_\_\_\_\_\_\_\_\_\_ (from Table F).

**Step 4:**

The test statistic is \_\_\_\_\_\_\_\_ extreme than the critical value, so the decision is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.

Or

Using d.f. = 14, the test statistic is larger than \_\_\_\_\_\_\_\_\_\_\_\_, so the *P*-value is \_\_\_\_\_\_\_\_\_ than 0.005, which also indicates that the null hypothesis is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ at the 0.01 level. (The *P*-value, from a calculator or computer, is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.)

**Step 5:**

# 8 – 4. *z* Test for a Proportion

## Objective 7. Test proportions Using the *z* Test.

Hypothesis tests involving proportions can be considered as a binomial experiment when there are only \_\_\_\_\_\_ outcomes and the probability of success does not \_\_\_\_\_\_\_\_\_\_ from trial to trial. Recall that the mean is and the standard deviation is for the binomial distribution. A normal distribution can be used to approximate the binomial distribution when and , the standard normal distribution can be used to test hypotheses for proportions.

### Test Statistic for the *z* Test for Proportions

where is the \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_,

is the population proportion, and

is the \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_.

The formula is derived from by substituting and and dividing numerator and denominator by .

### Assumptions for Testing a Proportion

1. The sample is a random sample.
2. The conditions for a binomial experiment are satisfied.
3. and .

### Example 8 − 9. Female Physicians

The percentage of physicians who are women is 27.9%. In a survey of physicians employed by a large university health system, 45 of 120 randomly selected physicians were women. Is there sufficient evidence at the 0.05 level of significance to conclude that the proportion of women physicians at the university health system exceeds 27.9%?

*Solution:*

*H*0: *p* = 27.9% or 0.279

The percentage of physicians who are women is 27.9%.

*H*1: *p* > 27.9% or 0.279

The proportion of women physicians at the university health system is larger than 27.9%. This is the research claim.

The test statistic is , rounded to two decimal places.

The critical value of *z* is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

For the right tailed test, the test statistic is to the right of the critical value, indicating to \_\_\_\_\_\_\_\_\_\_\_\_\_\_ the null hypothesis.

The *P*-value is \_\_\_\_\_\_\_\_\_\_. .

There i\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_the claim that the proportion of women physicians at the university health system exceeds 27.9%.

### Example 8 − 10. Exercise to Reduce Stress

A survey by *Men’s Health* magazine stated that 14% of men said they used exercise to reduce stress. Use . A random sample of 100 men was selected and 12 said that they exercise to relieve stress. Use the *P*-value method to test the claim. Is it reasonable to generalize the results to all adult Americans?

*Solution:*

*H*0: *p* = 14% or 0.14 This is the research claim.

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

*H*1: *\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

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

The test statistic is , rounded to two decimal places.

The critical value of *z* is .

The area on the *z* table associated with is \_\_\_\_\_\_. This is a two-tailed test, so double the area. Thus, .

There is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ the claim that the proportion of men who say they exercise to reduce stress is 14%.

It (is / is not) reasonable to generalize the results to all adult Americans, because the population from which the sample is selected is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

# 8 – 5. χ2 Test for a Variance or Standard Deviation

## Objective 8. Test Variances or Standard Deviations, Using the Chi-Square Test.

We used the chi-square distribution to construct confidence intervals for a single variance or standard deviation. Now, we will use the distribution to test a claim about a single variance or standard deviation.

### Characteristics of the Chi-Square Distribution

1. All chi-squared 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.

We will find the area under the chi-square distribution, using Table G in Appendix A, for right-tailed, left-tailed and two-tailed hypotheses.

Table G gives area to the \_\_\_\_\_\_\_\_\_\_\_ of the critical value.

For a right-tailed test, use the area under the significance level, *α*, for the specific degrees of freedom.

For a left-tailed test, subtract and use the area in the table for that value for a specific degrees of freedom.

For a two-tailed test, divide *α* by 2 and use the area under that value for a specific degrees of freedom for the right critical value and the area for value for the degrees of freedom for the left critical value.

Excerpt of Table G for the Chi Square Distribuion.  The left margin lists degrees of freedom from 1 to 30, that is, for sample sizes from 2 to 29.  The row at the top of the table shows the value of the area of the right tail, called alpha.  The values are 0.995, 0.99, 0.975, 0.95, 0.90, 0.10, 0.05, 0.025, 0.01, and 0.005.  The body of the table gives the critical values of chi square for each degrees of freedom and alpha.

### Example 8 − 11. Find Critical Values of the Chi-Square Distribution

Find the critical value for *χ*2:

1. Right-tailed test with 15 degrees of freedom and .
2. Left-tailed test with 27 degrees of freedom and .
3. Two-tailed test with 9 degrees of freedom and .

*Solution:*

1. The test is right-tailed, so we find the Chi-Square value on the 15th row representing 15 degrees of freedom and the column labeled . The critical value is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
2. The test is left-tailed, so find . Find the Chi-Square value on the 27th row for 27 degrees of freedom in the column labeled . The critical value is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
3. The test is two-tailed, so find and . Find the Chi-Square value on the 9th row for 9 degrees of freedom in the columns labeled \_\_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_\_\_. The critical values are \_\_\_\_\_\_\_\_\_\_and \_\_\_\_\_\_\_\_\_\_\_\_\_.

### Formula for Chi-Square Test Statistic for a Single Variance

with degrees of freedom equal to and where

*n* = sample size

*s2* = sample variance

*σ2* = population variance

### Assumptions for the Chi-Square Test for a Single Variance

1. The sample must be randomly selected from the population.
2. The population must be normally distributed for the variable under study.
3. The observations must be independent of one another.

### Example 8 − 12. Outpatient surgery (Use the Traditional Method.)

A hospital administrator believes that the standard deviation of the number of people using outpatient surgery per day is greater than 8. A random sample of 15 days is selected. The data are

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 25 | 30 | 5 | 15 | 18 | 42 | 16 | 9 |
| 10 | 12 | 12 | 38 | 8 | 14 | 27 |  |

At , is there enough evidence to support the administrator’s claim? Assume the variable is normally distributed.

*Solution:*

**Step 1: State the hypotheses and identify the claim.**

*H*0: . The standard deviation of the number of people using outpatient surgery per day is 8.

*H*1: . The standard deviation of the number of people using outpatient surgery per day is greater than 8. (This is the claim. It is right-tailed.)

**Step 2: Identify the critical value.**

**Step 3: Compute the test value.**

=

**Step 4: Make the decision.**

Choose one: Reject the null hypothesis OR Do not reject the null hypothesis.

Why?

**Step 5: Summarize the results.**

Table G can also be used to determine an interval for the P-values. Remember that the chi-square distribution is not symmetric and values cannot be negative.

### Example 8 − 13. Approximate the *P*-values

Find the *P*-value for :

* 1. Right-tailed test where and .
  2. Left-tailed test where and .
  3. Two-tailed test where and .

*Solution:*

1. The test is right-tailed with 8 degrees of freedom. In Table G, on the row labeled 8, find the values of that are smaller and larger than 19.274 that is, \_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_, corresponding to the columns labeled \_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_, respectively. Thus the *P*-value is between \_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_\_. \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_ (The calculated *P*-value is \_\_\_\_\_\_\_\_\_\_\_\_.)
2. The test is left-tailed with 22 degrees of freedom. In Table G, on the row labeled 22, notice that the first value is larger than 3.823. The heading for the first column is \_\_\_\_\_\_\_\_\_. Since this is a left-tailed test, subtract 1− α = \_\_\_\_\_\_\_\_\_\_\_\_\_\_. Then the *P*-value is less than \_\_\_\_\_\_\_\_\_\_\_. (The calculated P-value is \_\_\_\_\_\_\_\_\_\_\_\_\_\_.)
3. The test is two-tailed with 23 degrees of freedom. In Table G, on the row labeled 23 degrees of freedom, falls between \_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_, corresponding to \_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_, respectively. When the test value falls on the left side of the table, each of the values must be subtracted from one. \_\_\_\_\_ = \_\_\_\_\_\_ and \_\_\_\_\_\_ = \_\_\_\_\_. Then, since the test is two-tailed, each must be doubled to find the interval containing the *P*-value. Thus, \_\_\_\_\_\_ \_\_\_\_\_\_\_. (The calculated *P*-value is \_\_\_\_\_\_\_\_\_\_\_.)

### Example 8 − 14. Heights of Volcanos (Use the *P*-value Method.)

A random sample of heights (in feet) of active volcanoes in North America, outside of Alaska, follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 10,777 | 8159 | 11,240 | 10,456 | 14,163 | 8363 |

Is there sufficient evidence that the standard deviation in heights of volcanoes outside Alaska is less than the standard deviation in heights of Alaskan volcanoes, which is 2385.9 feet? Use

*Solution:*

**Step 1: State the hypotheses and identify the claim.**

**Step 2: Compute the test value.**

=

**Step 3: Find the *P*-value.**

**Step 4: Make the decision.**

Choose one: Reject the null hypothesis OR Do not reject the null hypothesis.

Why?

**Step 5: Summarize the results.**

# 8 – 6. Additional Topics Regarding Hypothesis Testing

## Objective 9. Test hypotheses, Using Confidence Intervals.

There is a relationship between confidence intervals and hypothesis testing. When the null hypothesis is rejected in a hypothesis testing situation, the confidence interval for the mean using the same level of significance **(will / will not)** contain the hypothesized mean. Also, when the null hypothesis is not rejected in a hypothesis testing situation, the confidence interval for the mean using the same level of significance **(will / will not)** contain the hypothesized mean.

### Example 8 − 15. One-Way Airfares

The average one-way airfare from Pittsburgh to Washington, D.C., is $236. A random sample of 20 one-way fares during a particular month had a mean of $210 with a standard deviation of $43. At , is there sufficient evidence to conclude a difference from the stated mean? Use the sample statistics to construct a 98% confidence interval for the true mean one-way airfare from Pittsburgh to Washington, D.C., and compare your interval to the results of the test. Do they support or contradict one another?

*Solution:*

**Step 1:**

**Step 2:**

**Step 3:**

**Step 4:**

**Step 5:**

The confidence interval for the mean is given by

The 98% confidence interval (does / does not) contain the hypothesized mean. Therefore, there (is / is not) agreement between the result of the hypothesis test and the confidence interval.

## Objective 10. Explain the Relationship Between Type I and Type II Errors and the Power of a Test.

Recall that in hypothesis testing, there are four possibilities:

The null hypothesis is rejected, and is actually \_\_\_\_\_\_\_\_\_. (Correct decision)

The null hypothesis is not rejected, and is actually \_\_\_\_\_\_\_\_\_. (Correct decision)

The null hypothesis is rejected, but is actually \_\_\_\_\_\_\_\_\_. (Type I error)

The null hypothesis is not rejected, but is actually \_\_\_\_\_\_\_\_\_. (Type II error)

A Type I error can only occur when the null hypothesis is \_\_\_\_\_\_\_\_\_\_\_\_\_\_. Remember that, when the level significance is chosen, researchers determine the maximum probability of making a Type I error.

A Type II error can only occur when the null hypothesis is ­­­­­­­­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. The probability of making a Type II error, called *β*, is not easy to calculate, but it is affected by several things, such as the value of \_\_\_\_\_\_\_\_ the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, the population standard deviation, and the difference between the claimed parameter and the true parameter. Researchers can only control the value of \_\_\_\_\_\_\_\_ and the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. Sometimes the population standard deviation can be estimated. The unknown quantity is the difference between the claimed parameter and the true parameter. If this difference were known, there would be no need to do hypothesis testing because we would know the true parameter.

We should not ignore the probability of making a Type II error, *β*, so researchers attempt to \_\_\_\_\_\_\_\_\_\_\_\_\_\_ *β* or to \_\_\_\_\_\_\_\_\_\_\_\_\_\_ , called the power of a hypothesis test. The power of a statistical test measures the sensitivity of the test to detect a real difference in parameters, if one actually exists. The higher the power of the test, the better the test is for rejecting the null hypothesis when it is actually false. For instance, if it were somehow known that *β* is 0.03, then the power is \_\_\_\_\_\_\_\_\_\_\_\_\_, meaning that the probability of rejecting a null hypothesis that is actually false is \_\_\_\_\_\_\_%.

If more than one test is appropriate, the test having the \_\_\_\_\_\_\_\_\_\_ power, as long as the test’s assumptions are met, should be used.

To increase the power of a test, a researcher can \_\_\_\_\_\_\_\_\_\_\_ the sample size or \_\_\_\_\_\_\_\_ the significance level. However, these methods have consequences. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_the significance level, increases the likelihood of committing a Type 1 error. \_\_\_\_\_\_\_\_\_\_\_\_\_ the sample size may increase the cost of the study or the time needed to organize the data. There are other methods to increase the power of a test that are beyond the scope of this course.

When the results of a statistical test indicate that a null hypothesis is not rejected, it may be that the null hypothesis is actually \_\_\_\_\_\_\_\_\_\_\_, but the power of the test is too low. Thus, a decision to not reject a null hypothesis does not mean that there is not enough evidence to support the alternative hypothesis, just that there is not enough evidence to reject the null hypothesis.

Our goal is to keep the probability of committing Type I or Type II errors as \_\_\_\_\_\_\_\_as possible.