9| Hypothesis Testing with One Sample

Chapter 9 Table of Contents

9| Hypothesis Testing with One Sample 1
Introduction 2
9.1 | Null and Alternative Hypotheses 4
9.2 | Outcomes and the Type 1 and 2 Errors 6
9.3 | Testing the Hypotheses 9
Hypothesis Test for the Mean 9
P-Value Approach 10
One and Two-tailed Tests 12
A Systematic Approach for Testing A Hypothesis 13
9.4 | Full Hypothesis Test Examples 14
Chapter 9 Try It Solutions 22
KEY TERMS 24
CHAPTER REVIEW 25

Adapted from: Claude Laflamme, Business Statistics -- BSTA 200 -- Humber College -- Version 2016 Revision A. OpenStax CNX. Jan 10, 2020 http://cnx.org/contents/0bec2053-5fc5-49c9-a97a-c33b7e0a095c@12.342.

Introduction

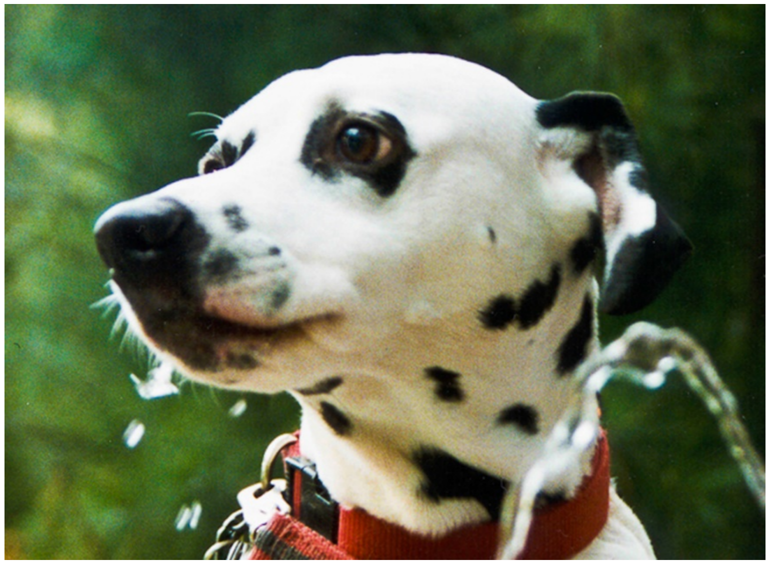


Figure 9.1 *You can use a hypothesis test to decide if a dog breeder's claim that every Dalmatian has 35 spots is statistically sound. (Credit: Robert Neff)*

Now we are down to the bread and butter work of the statistician: developing and testing hypotheses. It is important to put this material in a broader context so that the method by which a hypothesis is formed is understood completely. Using textbook examples often clouds the real source of statistical hypotheses.

Statistical testing is part of a much larger process known as the scientific method. This method was developed more than two centuries ago as the accepted way that new knowledge could be created. Until then, and unfortunately even today, among some, "knowledge" could be created simply by some authority saying something was so. Superstition and conspiracy theories were (are?) accepted uncritically.

The scientific method, briefly, states that only by following a careful and specific process can some assertion be included in the accepted body of knowledge. This process begins with a set of assumptions upon which a theory, sometimes called a model, is built. This theory, if it has any validity, will lead to predictions; what we call hypotheses.

As an example, in Microeconomics the theory of consumer choice begins with certain assumption concerning human behavior. From these assumptions a theory of how consumers make choices using indifference curves and the budget line. This theory gave rise to a very important prediction, namely, that there was an inverse relationship between price and quantity demanded. This relationship was known as the demand curve. The negative slope of the demand curve is really just a prediction, or a hypothesis, that can be tested with statistical tools.

Unless hundreds and hundreds of statistical tests of this hypothesis had not confirmed this relationship, the so-called Law of Demand would have been discarded years ago. This is the role of statistics, to test the hypotheses of various theories to determine if they should be admitted into the accepted body of knowledge; how we understand our world. Once admitted, however, they may be later discarded if new theories come along that make better predictions.

Not long ago two scientists claimed that they could get more energy out of a process than was put in. This caused a tremendous stir for obvious reasons. They were on the cover of *Time* and were offered extravagant sums to bring their research work to private industry and any number of universities. It was not long until their work was subjected to the rigorous tests of the scientific method and found to be a failure. No other lab could replicate their findings. Consequently, they have sunk into obscurity and their theory discarded. It may surface again when someone can pass the tests of the hypotheses required by the scientific method, but until then it is just a curiosity. Many pure frauds have been attempted over time, but most have been found out by applying the process of the scientific method.

This discussion is meant to show just where in this process statistics falls. Statistics and statisticians are not necessarily in the business of developing theories, but in the business of testing others' theories. Hypotheses come from these theories based upon an explicit set of assumptions and sound logic. The hypothesis comes first, before any data are gathered. Data do not create hypotheses; they are used to test them. If we bear this in mind as we study this section the process of forming and testing hypotheses will make more sense.

One job of a statistician is to make statistical inferences about populations based on samples taken from the population. **Confidence intervals** are one way to estimate a population parameter. Another way to make a statistical inference is to make a decision about the value of a specific parameter. For instance, a car dealer advertises that its new small truck gets 35 miles per gallon, on average. A tutoring service claims that its method of tutoring helps 90% of its students get an A or a B. A company says that women managers in their company earn an average of $60,000 per year.

A statistician will make a decision about these claims. This process is called "**hypothesis testing**." A hypothesis test involves collecting data from a sample and evaluating the data. Then, the statistician makes a decision as to whether or not there is sufficient evidence, based upon analyses of the data, to reject the null hypothesis.

In this chapter, you will conduct hypothesis tests on single means and single proportions. You will also learn about the errors associated with these tests.

9.1 | Null and Alternative Hypotheses

The actual test begins by considering two **hypotheses**. They are called the **null hypothesis** and the **alternative hypothesis**. These hypotheses contain opposing viewpoints.

H_0: **The null hypothesis:** It is a statement a statement about the population parameter that is the commonly accepted fact. This can often be considered the status quo.

H_1: **The alternative hypothesis:** It is a claim about the population that is contradictory to H_0 and what we conclude when we cannot accept H_0. The alternative hypothesis is the contender and must win with significant evidence to overthrow the status quo.

Since the null and alternative hypotheses are contradictory, you must examine evidence to decide if you have enough evidence to reject the null hypothesis or not. The evidence is in the form of sample data.

After you have determined which hypothesis the sample supports, you make a **decision.** There are two options for a decision. They are "**reject H_o**" if the sample information favors the alternative hypothesis or "**do not reject H_o**" if the sample information is insufficient to reject the null hypothesis. These conclusions are all based upon a level of probability, a significance level, that is set by the analyst.

Table 9-1 presents the various hypotheses in the relevant pairs. For example, if the null hypothesis is equal to some value, the alternative has to be not equal to that value.

H_0 **(Null Hypothesis)**	H_1 **(Alternative Hypothesis)**
equal (=)	not equal (≠)
greater than or equal to (≥)	less than (<)
less than or equal to (≤)	more than (>)

Table 9-1

NOTE
As a mathematical convention H_0 always has a symbol with an equal in it. H_1 never has a symbol with an equal in it. The choice of symbol depends on the wording of the hypothesis test.

Example 9-1

Suppose we want to test whether the mean GPA of students in American colleges is different from 2.0 (out of 4.0). Write the null and alternative hypotheses.

Solution 9-1

We start with the assumption that the mean GPA **is equal** to 2.0. This gives us our null hypothesis. The alternative hypothesis is that the mean GPA **is not equal** to 2.0.
Therefore, the null and alternative hypotheses are:

$$H_0: \mu = 2.0$$
$$H_1: \mu \neq 2.0$$

Example 9-2

We want to test if college students take less than five years to graduate from college, on the average. Write the null and alternative hypotheses.

Solution 9-2

We start with the assumption that students take **at least** five years, on average, to graduate. This is our null hypothesis. The alternative hypothesis is that students take **less than** five years to graduate, on average.

The null and alternative hypotheses are:

$$H_0: \mu \geq 5$$
$$H_1: \mu < 5$$

Try It 9-1

The mean entry level salary of an employee at a company is $58,000. You believe it is higher for IT professionals in the company. State the null and alternative hypotheses.

Chapter 9 Try It Solutions

Try It 9-2

Suppose that a recent article stated that the mean time spent in jail by a first–time convicted burglar has historically been 2.5 years. A study was then done to see if the mean time has changed in the new century. If you were conducting a hypothesis test to determine if the mean length of jail time has changed, what would the null and alternative hypotheses be?

Chapter 9 Try It Solutions

Try It 9-3

Your cable company claims that the mean speed of your cable Internet connection is at least three Megabits per second. You believe the mean speed is less than three Megabits per second. If you would like to test this claim, state the null and alternative hypotheses.

Chapter 9 Try It Solutions

9.2 | Outcomes and the Type 1 and 2 Errors

When you perform a hypothesis test, there are four possible outcomes depending on the actual truth (or falseness) of the null hypothesis H_0 and the decision to reject or not. The outcomes are summarized in the following table:

	H_0 is actually...	
Statistical Decision	True	False
Do not reject H_0	Correct outcome	Type II error
Reject H_0	Type I error	Correct outcome

The four possible outcomes in the table are:

1. The decision is **do not reject H_o** when **H_o is true (correct decision).**
2. The decision is **reject H_o** when **H_o is true** (incorrect decision known as a **Type I error**). This case is described as "rejecting a good null". As we will see later, it is this type of error that we will guard against by setting the probability of making such an error. The goal is to NOT take an action that is an error.
3. The decision is **do not reject H_o** when, in fact, **H_o is false** (incorrect decision known as a **Type II error**). This is called "accepting a false null". In this situation you have allowed the status quo to remain in force when it should be overturned. As we will see, the null hypothesis has the advantage in competition with the alternative.
4. The decision is **reject H_o** when **H_o is false** (**correct decision**).

Each of the errors occurs with a particular probability. The Greek letters α and β represent the probabilities.

α = probability of a Type I error = *P*(Type I error) = probability of rejecting the null hypothesis when the null hypothesis is true: rejecting a good null.

β = probability of a Type II error = *P*(Type II error) = probability of not rejecting the null hypothesis when the null hypothesis is false. **(1 − β)** is called the **Power of the Test**.

α and β should be as small as possible because they are probabilities of errors.

Statistics allows us to set the probability that we are making a Type I error. The probability of making a Type I error is α. Recall that the confidence intervals in the last unit were set by choosing an alpha value that determined the confidence level of the estimate because it was the probability of the interval failing to capture the true population parameter. This alpha and that one are the same.

The following are examples of Type I and Type II errors.

Example 9-3

Suppose the null hypothesis, H_0, is: Frank's rock climbing equipment is safe. The alternative hypothesis, H_1, is: Frank’s rock climbing equipment is not safe. State the possible Type 1 and Type 2 errors and explain the associated probabilities. Which type of error has the greater consequence?

Solution 9-3

Type I error: Frank thinks that his rock climbing equipment may not be safe when, in fact, it really is safe.

Type II error: Frank thinks that his rock climbing equipment may be safe when, in fact, it is not safe.

α = probability that Frank thinks his rock climbing equipment may not be safe when, in fact, it really is safe. **β = probability** that Frank thinks his rock climbing equipment may be safe when, in fact, it is not safe.

Notice that, in this case, the error with the greater consequence is the Type II error. (If Frank thinks his rock climbing equipment is safe, he will go ahead and use it.)

This is a situation described as "accepting a false null".

Example 9-4

Suppose the null hypothesis, H_0, is: The victim of an automobile accident is alive when he arrives at the emergency room of a hospital. State the possible Type 1 and Type 2 errors and explain the associated probabilities. Which type of error has the greater consequence?

Solution 9-4

Type I error: The emergency crew thinks that the victim is dead when, in fact, the victim is alive. **Type II error**: The emergency crew does not know if the victim is alive when, in fact, the victim is dead.

α = probability that the emergency crew thinks the victim is dead when, in fact, he is really alive = P(Type I error). **β = probability** that the emergency crew does not know if the victim is alive when, in fact, the victim is dead = P(Type II error).

The error with the greater consequence is the Type I error. (If the emergency crew thinks the victim is dead, they will not treat him.)

Try It 9-4

Suppose the null hypothesis, H_0, is: a patient is not sick. Which type of error has the greater consequence, Type I or Type II?

Chapter 9 Try It Solutions

Try It 9-5

"Red tide" is a bloom of poison-producing algae–a few different species of a class of plankton called dinoflagellates. When the weather and water conditions cause these blooms, shellfish such as clams living in the area develop dangerous levels of a paralysis-inducing toxin. In Massachusetts, the Division of Marine Fisheries (DMF) monitors levels of the toxin in shellfish by regular sampling of shellfish along the coastline. If the mean level of toxin in clams exceeds 800 μg (micrograms) of toxin per kg of clam meat in any area, clam harvesting is banned there until the bloom is over and levels of toxin in clams subside. Describe both a Type I and a Type II error in this context, and state which error has the greater consequence.

Chapter 9 Try It Solutions

9.3 | Testing the Hypotheses

Earlier, we discussed sampling distributions. Particular distributions are associated with hypothesis testing. We will perform **hypotheses tests of a population mean** using a **normal distribution** or a **Student's t-distribution**. (Remember, use a Student's t-distribution when the population **standard deviation is unknown**.)

Hypothesis Test for the Mean

When the population standard deviation $\boldsymbol{\sigma}$ **is known** we can derive the **test statistic** for testing hypotheses concerning means as follows:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

When the population standard deviation $\boldsymbol{\sigma}$ **is unknown** we can derive the **test statistic** for testing hypotheses concerning means as follows:

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

$df = n - 1$ (population is assumed to be normal for small n)

The standardizing formula cannot be solved as it is because we do not have μ, the population mean. However, if we substitute the value of the mean given in the null hypothesis into one of the formulas as above, we can compute a z or t value. We interpret this z value as the associated probability that a sample with a sample mean of $\bar{x}$ could have come from a distribution with a population mean equal to the value in the null hypothesis H_0.

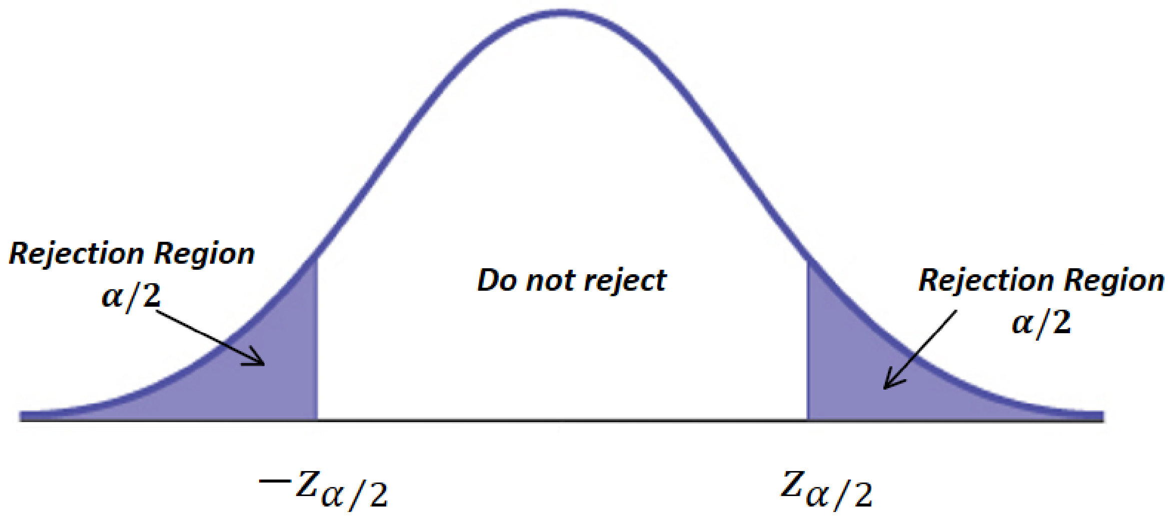


Figure 9.2

In Figure 9.2 we see a graphical representation of the decision-making process for a two tailed test. $z_{\alpha/2}$ and $-z_{\alpha/2}$, called the **critical values**, are marked on the bottom panel as the z values associated with the probability the analyst has set as the level of significance in the test, (α). The probabilities in the tails of both panels are equal to $\alpha/2$ so that the α value is split evenly between the two tails. If the calculated test statistic falls into the rejection region, that is less than $-z_{\alpha/2}$ or greater than $z_{\alpha/2}$ the null hypothesis, H_0, will be rejected. If the test statistic does not fall into the rejection region, the null hypothesis will not be rejected.

This rule will always be the same no matter what hypothesis we are testing or what formulas we are using to make the test. Stating the decision rule another way: if the sample mean is unlikely to have come from the distribution with the hypothesized mean we cannot accept the null hypothesis. Here we define "unlikely" as having a probability less than alpha of occurring.

P-Value Approach

An alternative decision rule can be developed by calculating the probability that a sample mean could be found that would give a test statistic larger than the test statistic found from the current sample data assuming that the null hypothesis is true. Here the notion of "likely" and "unlikely" is defined by the probability of drawing a sample with a mean from a population with the hypothesized mean that is either larger or smaller than that found in the sample data. Simply stated, the p-value approach compares the desired significance level, α, to the p-value which is the probability of drawing a sample mean further from the hypothesized value than the actual sample mean. A **large *p*-value** calculated from the data indicates that we should **not reject the null hypothesis**. The **smaller the *p*-value**, the more unlikely the outcome, and **the stronger the evidence is against the null hypothesis**. We would reject the null hypothesis if the evidence is strongly against it.

If we use the *p*-value decision rule we need one more step. Using software or our tables, we need to find the probability associated with the calculated test statistic. We then compare that to the α associated with our selected level of confidence.

Here is a systematic way to make a decision of whether you cannot accept or cannot reject a null **hypothesis** if using the ***p*-value** and a **preset or preconceived α** (the "**significance level**"). A preset α is the probability of a **Type I** error (rejecting the null hypothesis when the null hypothesis is true). It may or may not be given to you at the beginning of the problem. In any case, the value of α is the decision of the analyst. When you make a decision to reject or not reject H_o, do as follows:

- If α > *p*-value, we reject H_o. The results of the sample data are significant. There is sufficient evidence to conclude that H_0 is an incorrect belief and that the **alternative hypothesis**, H_1, may be correct.
- If $\alpha \leq$ *p*-value, we do not reject H_o. The results of the sample data are not significant. There is not sufficient evidence to conclude that the alternative hypothesis, H_1, may be correct. In this case the status quo stands.
- When you "do not reject H_o", it does not mean that you should believe that H_o is true. It simply means that the sample data have **failed** to provide sufficient evidence to cast serious doubt about the truthfulness of H_o. Remember that the null is the status quo and it takes high probability to overthrow the status quo. This bias in favor of the null hypothesis is what gives rise to the statement "tyranny of the status quo" when discussing hypothesis testing and the scientific method.

Both decision rules will result in the same decision and it is a matter of preference which one is used.

One and Two-tailed Tests

Figure 9.2 is a graphical representation of a **two-tailed test**. Two-tailed tests occur when the alternative hypothesis allows that the mean is **different from** the hypothesized mean – that is, it could have come from a population which was either larger or smaller than the hypothesized mean in the null hypothesis. With a two-tailed test, α is split into two and half of the α value goes into each tail.

A test of a claim that is one-sided will be a **one-tailed test**. For example, a car manufacturer claims that their Model 17B provides gas mileage of **greater than** 25 miles per gallon. The null and alternative hypothesis would be:

- $H_0: \mu \leq 25$
- $H_1: \mu > 25$

The claim would be in the alternative hypothesis. The burden of proof in hypothesis testing is carried in the alternative. This is because failing to reject the null, the status quo, must be accomplished with 90 or 95 percent significance that it cannot be maintained. Said another way, we want to have only a 5 or 10 percent probability of making a Type I error, rejecting a good null; overthrowing the status quo.

This is a one-tailed test and all of the alpha probability is placed in just one tail and not split into $\alpha/2$ as in the above case of a two-tailed test.

Figure 9.3 shows the two possible cases and the form of the null and alternative hypothesis that give rise to them.

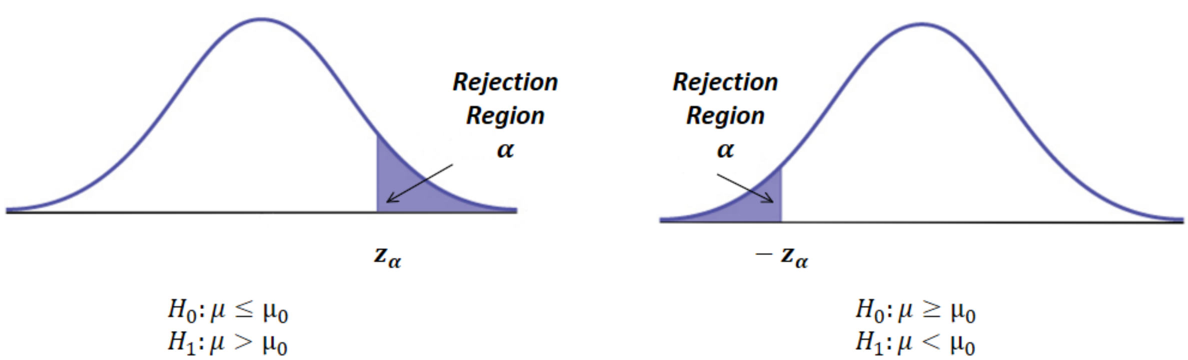


Figure 9.3 Graphical representation of one-tailed tests.

In Figure 9.3 μ_0 is the hypothesized value of the population mean.

A Systematic Approach for Testing A Hypothesis

A systematic approach to hypothesis testing follows the following steps and in this order. This template will work for all hypotheses that you will ever test.

- Set up the null and alternative hypothesis. This is typically the hardest part of the process. Here the question being asked is reviewed. What parameter is being tested, a mean, a proportion, differences in means, etc. Is this a one-tailed test or two-tailed test?
- Decide the level of significance required for this particular case and determine the critical value. These can be found in the appropriate statistical table. The levels of confidence typical for the social sciences are 90%, 95% and 99%. However, the level of significance is a policy decision and should be based upon the risk of making a Type I error, rejecting a good null. Consider the consequences of making a Type I error.
- Next, on the basis of the hypotheses and sample size, select the appropriate test statistic and find the relevant critical value: z, t etc. Drawing the relevant probability distribution and marking the critical value is always big help. Be sure to match the graph with the hypothesis, especially if it is a one-tailed test.
- Take a sample(s) and calculate the relevant parameters: sample mean, standard deviation, or proportion. Using the formula for the test statistic, now calculate the test statistic for this particular case using the parameters you have just calculated.
- Compare the calculated test statistic and the critical value. Marking these on the graph will give a good visual picture of the situation. There are now only two situations:
- The test statistic is in the tail: Reject the null hypothesis H_0, the probability that this sample mean (proportion) came from the hypothesized distribution is too small to believe that it is the real home of these sample data.
- The test statistic is not in the tail: Do not reject the null hypothesis H_0, the sample data are compatible with the hypothesized population parameter.
- Reach a conclusion. It is best to articulate the conclusion two different ways. First a formal statistical conclusion such as "With a 95 % level of significance we reject the null hypotheses that the population mean is equal to XX (units of measurement)". If the formal conclusion was

that above, then the informal one might be, “The machine is broken and we need to shut it down and call for repairs”.

All hypotheses tested will go through this same process. The only changes are the relevant formulas and those are determined by the hypothesis required to answer the original question.

Four-Step Hypothesis Testing

The process of hypothesis testing can be completed by following the following four steps.

Step 1: State the null and alternative hypotheses.

Step 2: Specify α and state the rejection region.

Step 3: Calculate the value of the test statistic.

Step 4: Make a decision and interpret the results.

These steps will be demonstrated with several examples.

9.4 | Full Hypothesis Test Examples

Example 9-5

Jeffrey, as an eight-year old, established a mean time of 16.43 seconds for swimming the 25-yard freestyle. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for 15 25-yard freestyle swims. For the 15 swims, Jeffrey's mean time was 16 seconds. Assume the population standard deviation is 0.8 seconds. Using $\alpha = 0.05$, conduct a hypothesis test to determine if Jeffrey swims faster with goggles.

Solution 9-5

Step 1: State the null and alternative hypotheses.

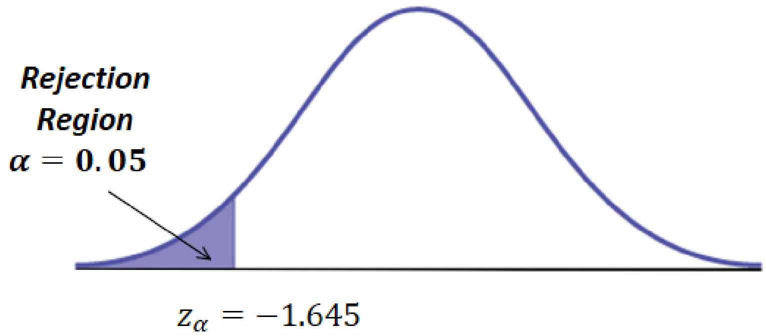
$$H_0: \mu \geq 16.43$$

$$H_1: \mu < 16.43$$

The alternative hypothesis, H_1, is testing the claim that Jeffrey swims faster with his goggles on; that is, less than his usual average of 16.43 seconds.

Step 2: Determine the rejection region.

Since we know the population standard deviation, σ, we will work with the z distribution. This is a one-tail, left-tailed test.

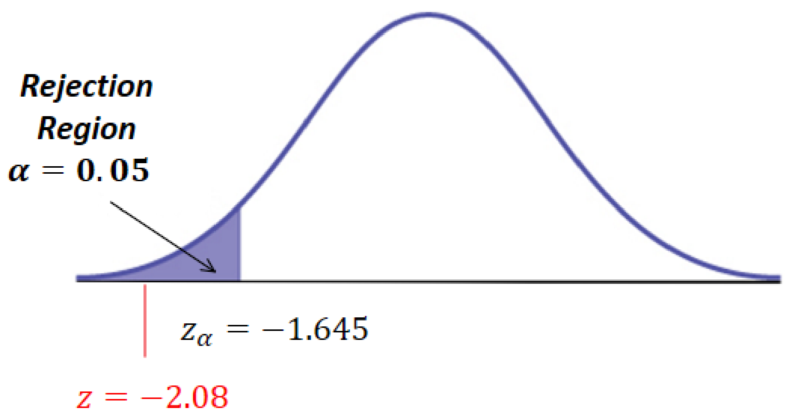


We will reject H_0 if the test statistic is less than -1.645.

Step 3: Calculate the test statistic.

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{16 - 16.43}{\frac{0.8}{\sqrt{15}}} = -2.08$$

Step 4: Make a decision and state your conclusion.



The test statistic $z = -2.08$ is less than the critical value $z = -1.645$. Therefore, we reject H_0. There is enough evidence to conclude that Jeffrey's swims faster, on average, with the goggles.

Example 9-6

The mean throwing distance of a football for Marco, a high school freshman quarterback, is 40 yards, with a standard deviation of two yards. The team coach tells Marco to adjust his grip to get more distance. The coach records the distances for 20 throws. For the 20 throws, Marco's mean distance was 45 yards with a standard deviation of two yards. The coach thought the different grip helped Marco throw farther than 40 yards. Conduct a hypothesis test using α = 0.10. Assume the throw distances for footballs are normal.

Solution 9-6

Step 1: State the null and alternative hypotheses.

$$H_0: \mu \leq 40$$

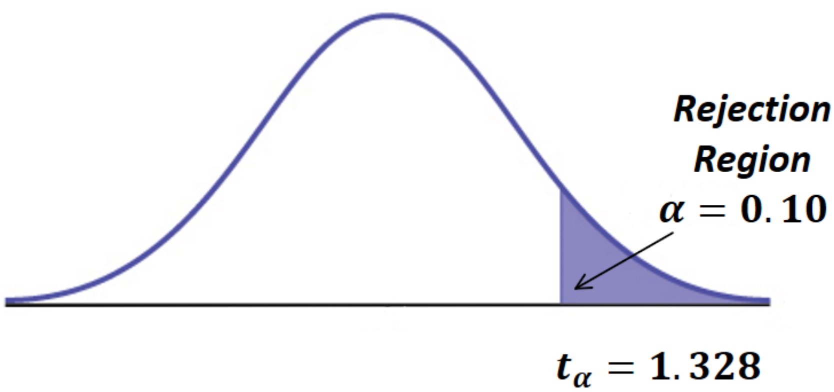
$$H_1: \mu > 40$$

The alternative hypothesis, H_1, is testing the claim that, with a different grip, Marco throws further than 40 yards.

Step 2: Determine the rejection region.

We do not know the population standard deviation, σ, so we will work with the t distribution. This is a one-tail, right-tailed test.

$$\alpha = 10\%, \quad df = 20 - 1 = 19, \qquad t_\alpha = 1.328$$

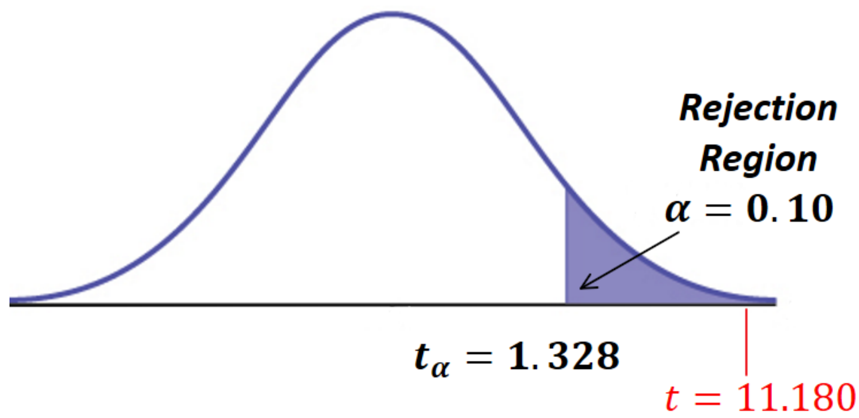


Reject H_0 if the test statistic is greater than 1.328.

Step 3: Calculate the test statistic.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{45 - 40}{\frac{2}{\sqrt{20}}} = 11.180$$

Step 4: Make a decision and state your conclusion.



The test statistic $t = 11.180$ is greater than the critical value $t = 1.328$. Therefore, we reject H_0. There is enough evidence to conclude that Marco throws further than 40 yards with the different grip.

Example 9-7

Jane has just begun her new job as on the sales force of a very competitive company. Salespeople at Jane's company close contracts for an average value of \$100. In a sample of 16 sales calls it was found that Jane closed the contract for an average value of \$104 with a standard deviation of \$12. Test at 5% level significance that the population mean of Jane's sales calls are different from \$100. Assume Jane's sales calls follow a normal distribution.

Solution 9-7

Step 1: State the null and alternative hypotheses.

$$H_0: \mu = 100$$

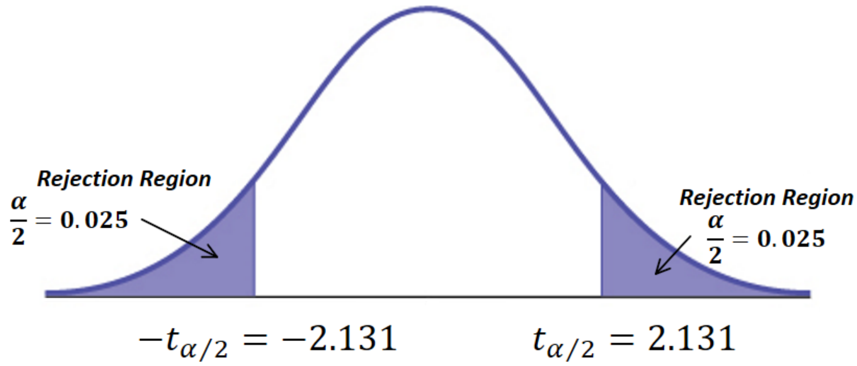
$$H_1: \mu \neq 100$$

The alternative hypothesis, H_1, is testing the claim that Jane's sales average is different from \$100.

Step 2: Determine the rejection region.

We do not know the population standard deviation, σ, so we will work with the t distribution. This is a two-tail test.

$$\alpha = 5\%, \quad df = 16 - 1 = 15, \qquad t_{\alpha/2} = \pm 2.131$$

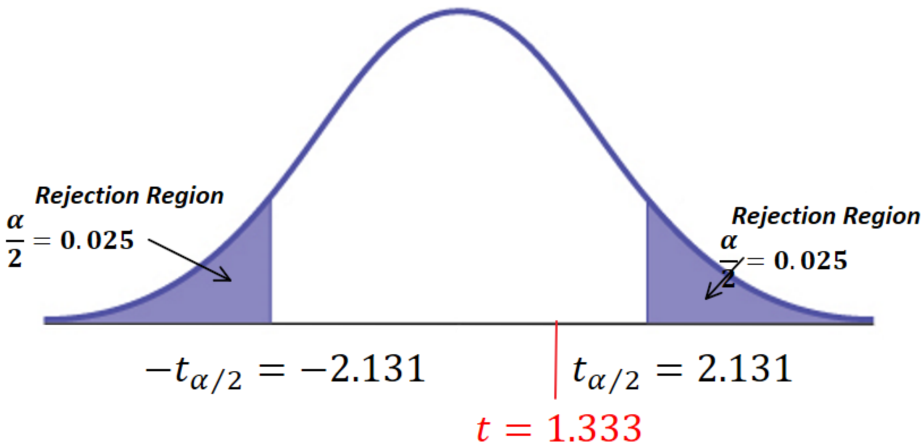


Reject H_0 if the test statistic is less than -2.131 or greater than 2.131.

Step 3: Calculate the test statistic.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{104 - 100}{\frac{12}{\sqrt{16}}} = 1.333$$

Step 4: Make a decision and state your conclusion.



The test statistic $t = 1.333$ does not fall in the rejection region. Therefore, we do not reject H_0. There is not enough evidence to conclude that Jane's sales are, on average, different from $100.

Example 9-8

It is believed that a stock price for a particular company will grow at a rate of $5 per week. An investor believes the stock won't grow as quickly. The changes in stock price is recorded for ten weeks and are as follows: $4, $3, $2, $3, $1, $7, $2, $1, $1, $2. Test the investor's claim using a 5% level of significance.

Solution 9-8

Pre-step: Calculate the sample mean and standard deviation.

In this example, we are given the sample data and must calculate the mean and standard deviation from this data. Using our formulas or software, we find:

$$\bar{x} = 2.60$$

$$s = 1.83787$$

Step 1: State the null and alternative hypotheses.

$$H_0: \mu \geq 5$$

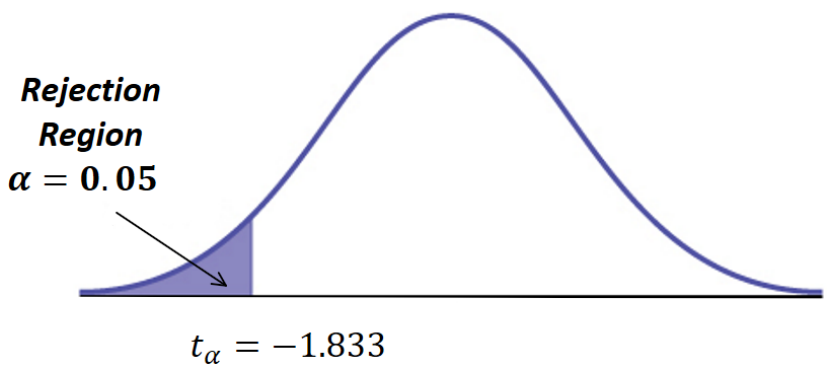
$$H_1: \mu < 5$$

The alternative hypothesis, H_1, is testing the claim the stock price grows slower than $5 per week.

Step 2: Determine the rejection region.

We do not know the population standard deviation, σ, so we will work with the t distribution. This is a one-tail, left-tail test.

$$\alpha = 5\%, \quad df = 10 - 1 = 9, \qquad t_\alpha = -1.833$$

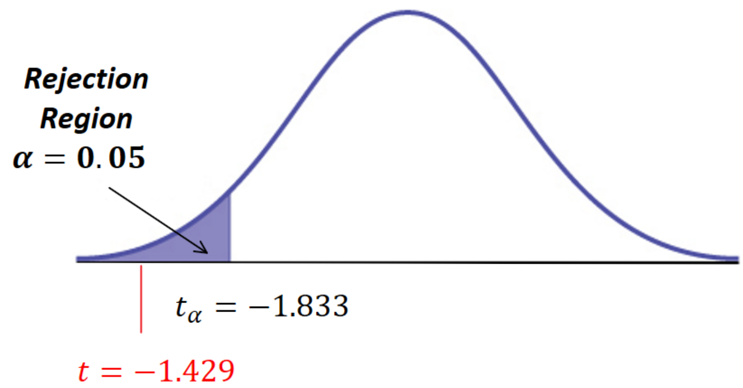


Reject H_0 if the test statistic is less than the critical value $t = -1.833$.

Step 3: Calculate the test statistic.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{2.60 - 5}{\frac{1.83787}{\sqrt{10}}} = -4.129$$

Step 4: Make a decision and state your conclusion.



The test statistic $t = -1.429$ is less than -1.833. Therefore, we reject H_0. There is enough evidence to conclude that the stock price, on average, is growing more slowly than $5 per week.

Try It 9-6

The package of your favorite cookies states that each cookie weighs 12 grams. You decide to test the accuracy of this claim. A sample of 7 randomly selected cookies yields a mean weight of 11.6 grams, with a standard deviation of 0.7 grams. At the 5% level of significance, can you conclude that the weight of each cookie is less than is stated on the package?

Chapter 9 Try It Solutions

Try It 9-7

A coffee machine is designed to fill cups to 250mL with a standard deviation of 5mL. A sample of 12 cups of coffee is taken. The mean of the sample is 246mL. Using $\alpha = 0.10$, can we conclude that the mean is not equal to 250mL?

Chapter 9 Try It Solutions

Try It 9-8

A high school states that the mean graduating average of their students is 72%, with a standard deviation of 12%. A random sample of 25 students is taken. The mean graduating average of the sample is 76%. Using $\alpha = 0.03$, can we conclude that the graduating average is more than 72%?

Chapter 9 Try It Solutions

Try It 9-9

You have recently purchased a new car. The manufacturer claims that the car has a fuel efficiency of 6.6 L/100 km during highway driving. Since you regularly drive long distances on the highway, you decide to test this claim. A sample of 8 test drives yields a mean of 7.1 L/100 km and a standard deviation of 0.5 L/100 km. At the 10% level of significance, can you conclude the fuel efficiency of your car during highway driving is different than the manufacturer's claim?

Chapter 9 Try It Solutions

Chapter 9 Try It Solutions

Try It 9-1

$H_0 : \mu \leq 58{,}000$

$H_1 : \mu > 58{,}000$

Try It 9-2

$H_0 : \mu = 2.5$

$H_1 : \mu \neq 2.5$

Try It 9-3

$H_0 : \mu \geq 3$

$H_1 : \mu < 3$

Try It 9-4

The error with the greater consequence is the Type II error: the patient will be thought well when, in fact, he is sick, so he will not get treatment.

Try It 9-5

In this scenario, an appropriate null hypothesis would be H_0: the mean level of toxins is at most 800 μg, H_0: $\mu_0 \leq 800\,\mu$g.

Type I error: The DMF believes that toxin levels are still too high when, in fact, toxin levels are at most 800 μg. The DMF continues the harvesting ban.

Type II error: The DMF believes that toxin levels are within acceptable levels (are at least 800 μg) when, in fact, toxin levels are still too high (more than 800 μg). The DMF lifts the harvesting ban. This error could be the most serious. If the ban is lifted and clams are still toxic, consumers could possibly eat tainted food.

In summary, the more dangerous error would be to commit a Type II error, because this error involves the availability of tainted clams for consumption.

Try It 9-6

$H_0 : \mu \geq 12,\ H_1 : \mu < 12$

Rejection region: Reject H_0 if $t < -1.943$

Test statistic: $t = -1.512$

Decision: Do not reject H_0. There is not enough evidence to conclude that the weight of a cookie is less than claimed.

Try It 9-7

$H_0: \mu = 250, \quad H_1: \mu \neq 250$

Rejection region: Reject H_0 if $z < -1.645$ or $z > 1.645$

Test statistic: $z = -2.77$

Decision: Reject H_0. There is enough evidence to conclude that the mean is not equal to 250 mL.

Try It 9-8

$H_0: \ \mu \leq 72, \ H_1: \ \mu > 72$

Rejection region: Reject H_0 if $z > 1.88$

Test statistic: $z = 1.67$

Decision: Do not reject H_0. There is not enough evidence to conclude that the graduating average is more than 72%.

Try It 9-9

$H_0: \mu = 6.6, \ H_1: \mu \neq 6.6$

Rejection region: Reject H_0 if $t < -1.895$ or $t > 1.895$.

Test statistic: $t = 2.828$

Decision: Reject H_0. There is enough evidence to conclude that the fuel efficiency is different than claimed.

KEY TERMS

Hypothesis Testing

Based on sample evidence, a procedure for determining whether the hypothesis stated is a reasonable statement and should not be rejected, or is unreasonable and should be rejected.

Hypothesis

a statement about the value of a population parameter, in case of two hypotheses, the statement assumed to be true is called the null hypothesis (notation H_0) and the contradictory statement is called the alternative hypothesis (notation H_1).

Level of Significance of the Test

probability of a Type I error (reject the null hypothesis when it is true). Notation: α. In hypothesis testing, the Level of Significance is called the preconceived α or the preset α. The Confidence level is (1-α).

CHAPTER REVIEW

In a **hypothesis test**, sample data is evaluated in order to arrive at a decision about some type of claim. If certain conditions about the sample are satisfied, then the claim can be evaluated for a population. In a hypothesis test, we:

1. Evaluate the **null hypothesis**, typically denoted with H_0. The null is not rejected unless the hypothesis test shows otherwise. The null statement must always contain some form of equality (=, ≤ or ≥)
2. Always write the **alternative hypothesis**, denoted with H_1, using not equal, less than or greater than symbols, i.e., (≠, <, or >).
3. If we reject the null hypothesis, then we can assume there is enough evidence to support the alternative hypothesis.
4. Never state that a claim is proven true or false. Keep in mind the underlying fact that hypothesis testing is based on probability laws; therefore, we can talk only in terms of non-absolute certainties.

In every hypothesis test, the outcomes are dependent on a correct interpretation of the data. Incorrect calculations or misunderstood summary statistics can yield errors that affect the results. A **Type I** error occurs when a true null hypothesis is rejected. A **Type II** error occurs when a false null hypothesis is not rejected.

The probabilities of these errors are denoted by the Greek letters α and β, for a Type I and a Type II error respectively. The power of the test, $1 - \beta$, quantifies the likelihood that a test will yield the correct result of a true alternative hypothesis being accepted. A high power is desirable.