Chapter 9

Hypothesis Testing: Single Mean and Single Proportion

9.1 Hypothesis Testing: Single Mean and Single Proportion[1]

9.1.1 Student Learning Objectives

By the end of this chapter, the student should be able to:

- Differentiate between Type I and Type II Errors
- Describe hypothesis testing in general and in practice
- Conduct and interpret hypothesis tests for a single population mean, population standard deviation known.
- Conduct and interpret hypothesis tests for a single population mean, population standard deviation unknown.
- Conduct and interpret hypothesis tests for a single population proportion.

9.1.2 Introduction

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 a parameter. For instance, a car dealer advertises that its new small truck gets 35 miles per gallon, on the 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 the data supports the claim that is made about the population.

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.

Hypothesis testing consists of two contradictory hypotheses or statements, a decision based on the data, and a conclusion. To perform a hypothesis test, a statistician will:

[1]This content is available online at <http://http://cnx.org/content/m16997/1.9/>.

1. Set up two contradictory hypotheses.
2. Collect sample data (in homework problems, the data or summary statistics will be given to you).
3. Determine the correct distribution to perform the hypothesis test.
4. Analyze sample data by performing the calculations that ultimately will support one of the hypotheses.
5. Make a decision and write a meaningful conclusion.

NOTE: To do the hypothesis test homework problems for this chapter and later chapters, make copies of the appropriate special solution sheets. See the Table of Contents topic "Solution Sheets".

9.2 Null and Alternate Hypotheses[2]

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

H_o: **The null hypothesis:** It is a statement about the population that will be assumed to be true unless it can be shown to be incorrect beyond a reasonable doubt.

H_a: **The alternate hypothesis:** It is a claim about the population that is contradictory to H_o and what we conclude when we reject H_o.

Example 9.1
H_o: No more than 30% of the registered voters in Santa Clara County voted in the primary election.

H_a: More than 30% of the registered voters in Santa Clara County voted in the primary election.

Example 9.2
We want to test whether the average grade point average in American colleges is 2.0 (out of 4.0) or not.

H_o: $\mu = 2.0$ H_a: $\mu \neq 2.0$

Example 9.3
We want to test if college students take less than five years to graduate from college, on the average.

H_o: $\mu \geq 5$ H_a: $\mu < 5$

Example 9.4
In an issue of **U. S. News and World Report**, an article on school standards stated that about half of all students in France, Germany, and Israel take advanced placement exams and a third pass. The same article stated that 6.6% of U. S. students take advanced placement exams and 4.4 % pass. Test if the percentage of U. S. students who take advanced placement exams is more than 6.6%.

H_o: $p = 0.066$ H_a: $p > 0.066$

Since the null and alternate hypotheses are contradictory, you must examine evidence to decide which hypothesis the evidence supports. The evidence is in the form of sample data. The sample might support either the null hypothesis or the alternate hypothesis but not both.

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 alternate hypothesis or

[2]This content is available online at <http://http://cnx.org/content/m16998/1.10/>.

"do not reject H_o" if the sample information favors the null hypothesis, meaning that there is not enough information to reject the null.

Mathematical Symbols Used in H_o and H_a:

H_o	H_a
equal (=)	not equal ($\neq$) **or** greater than (>) **or** less than (<)
greater than or equal to ($\geq$)	less than (<)
less than or equal to ($\leq$)	more than (>)

Table 9.1

NOTE: H_o always has a symbol with an equal in it. H_a never has a symbol with an equal in it. The choice of symbol depends on the wording of the hypothesis test. However, be aware that many researchers (including one of the co-authors in research work) use = in the Null Hypothesis, even with > or < as the symbol in the Alternate Hypothesis. This practice is acceptable because we only make the decision to reject or not reject the Null Hypothesis.

9.2.1 Optional Collaborative Classroom Activity

Bring to class a newspaper, some news magazines, and some Internet articles . In groups, find articles from which your group can write a null and alternate hypotheses. Discuss your hypotheses with the rest of the class.

9.3 Outcomes and the Type I and Type II Errors[3]

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

ACTION	H_o **IS ACTUALLY**	**...**
	True	False
Do not reject H_o	Correct Outcome	Type II error
Reject H_o	Type I Error	Correct Outcome

Table 9.2

The four outcomes in the table are:

- The decision is to **not reject** H_o when, in fact, H_o **is true (correct decision).**
- The decision is to **reject** H_o when, in fact, H_o **is true** (incorrect decision known as a **Type I error**).
- The decision is to **not reject** H_o when, in fact, H_o **is false** (incorrect decision known as a **Type II error**).
- The decision is to **reject** H_o when, in fact, H_o **is false** (**correct decision** whose probability is called the **Power of the Test**).

[3]This content is available online at <http://http://cnx.org/content/m17006/1.7/>.

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.

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

α and β should be as small as possible because they are probabilities of errors. They are rarely 0.

The Power of the Test is $1 - \beta$. Ideally, we want a high power that is as close to 1 as possible.

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

Example 9.5
Suppose the null hypothesis, H_o, is: Frank's rock climbing equipment is safe.

Type I error: Frank concludes that his rock climbing equipment may not be safe when, in fact, it really is safe. **Type II error**: Frank concludes that his rock climbing equipment is 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. β **= probability** that Frank thinks his rock climbing equipment is safe when, in fact, it is not.

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.)

Example 9.6
Suppose the null hypothesis, H_o, is: The victim of an automobile accident is alive when he arrives at the emergency room of a hospital.

Type I error: The emergency crew concludes that the victim is dead when, in fact, the victim is alive. **Type II error**: The emergency crew concludes that 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 thinks the victim is alive when, in fact, he 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.)

9.4 Distribution Needed for Hypothesis Testing[4]

Earlier in the course, we discussed sampling distributions. **Particular distributions are associated with hypothesis testing.** Perform tests of a population mean using a **normal distribution** or a **student-t distribution.** (Remember, use a student-t distribution when the population **standard deviation** is unknown and the population from which the sample is taken is normal.) In this chapter we perform tests of a population proportion using a normal distribution (usually n is large or the sample size is large).

If you are testing a **single population mean**, the distribution for the test is for **averages**:

[4]This content is available online at <http://http://cnx.org/content/m17017/1.10/>.

$\overline{X} \sim N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$ or t_{df}

The population parameter is μ. The estimated value (point estimate) for μ is $\overline{x}$, the sample mean.

If you are testing a **single population proportion**, the distribution for the test is for proportions or percentages:

$P' \sim N\left(p, \sqrt{\frac{p \cdot q}{n}}\right)$

The population parameter is p. The estimated value (point estimate) for p is p'. $p' = \frac{x}{n}$ where x is the number of successes and n is the sample size.

9.5 Assumption[5]

When you perform a **hypothesis test of a single population mean** μ using a **Student-t distribution** (often called a t-test), there are fundamental assumptions that need to be met in order for the test to work properly. Your data should be a **simple random sample** that comes from a population that is approximately **normally distributed**. You use the sample **standard deviation** to approximate the population standard deviation. (Note that if the sample size is larger than 30, a t-test will work even if the population is not approximately normally distributed).

When you perform a **hypothesis test of a single population mean** μ using a normal distribution (often called a z-test), you take a simple random sample from the population. The population you are testing is normally distributed or your sample size is larger than 30 or both. You know the value of the population standard deviation.

When you perform a **hypothesis test of a single population proportion** p, you take a simple random sample from the population. You must meet the conditions for a **binomial distribution** which are there are a certain number n of independent trials, the outcomes of any trial are success or failure, and each trial has the same probability of a success p. The shape of the binomial distribution needs to be similar to the shape of the normal distribution. To ensure this, the quantities np and nq must both be greater than five ($np > 5$ and $nq > 5$). Then the binomial distribution of sample (estimated) proportion can be approximated by the normal distribution with $\mu = p$ and $\sigma = \sqrt{\left(\frac{p \cdot q}{n}\right)}$. Remember that $q = 1 - p$.

9.6 Rare Events[6]

Suppose you make an assumption about a property of the population (this assumption is the **null hypothesis**). Then you gather sample data randomly. If the sample has properties that would be very **unlikely** to occur if the assumption is true, then you would conclude that your assumption about the population is probably incorrect. (Remember that your assumption is just an **assumption** - it is not a fact and it may or may not be true. But your sample data is real and it is showing you a fact that seems to contradict your assumption.)

For example, Didi and Ali are at a birthday party of a very wealthy friend. They hurry to be first in line to grab a prize from a tall basket that they cannot see inside because they will be blindfolded. There are 200 plastic bubbles in the basket and Didi and Ali have been told that there is only one with a $100 bill. Didi is the first person to reach into the basket and pull out a bubble. Her bubble contains a $100 bill. The

[5]This content is available online at <http://http://cnx.org/content/m17002/1.14/>.
[6]This content is available online at <http://http://cnx.org/content/m16994/1.6/>.

probability of this happening is $\frac{1}{200} = 0.005$. Because this is so unlikely, Ali is hoping that what the two of them were told is wrong and there are more $100 bills in the basket. A "rare event" has occurred (Didi getting the $100 bill) so Ali doubts the assumption about only one $100 bill being in the basket.

9.7 Using the Sample to Support One of the Hypotheses[7]

Use the sample (data) to calculate the actual probability of getting the test result, called the **p-value**. The p-value is the **probability that an outcome of the data (for example, the sample mean) will happen purely by chance when the null hypothesis is true**.

A large p-value calculated from the data indicates that the sample result is likely happening purely by chance. The data support the **null hypothesis** so we do not reject it. 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 the null hypothesis.

The p-value is sometimes called the **computed** α because it is calculated from the data. You can think of it as the probability of (incorrectly) rejecting the null hypothesis when the null hypothesis is actually true.

Draw a graph that shows the p-value. The hypothesis test is easier to perform if you use a graph because you see the problem more clearly.

Example 9.7: (to illustrate the p-value)
Suppose a baker claims that his bread height is more than 15 cm, on the average. Several of his customers do not believe him. To persuade his customers that he is right, the baker decides to do a hypothesis test. He bakes 10 loaves of bread. The average height of the sample loaves is 17 cm. The baker knows from baking hundreds of loaves of bread that the **standard deviation** for the height is 0.5 cm.

The null hypothesis could be H_o: $\mu \leq 15$ The alternate hypothesis is H_a: $\mu > 15$

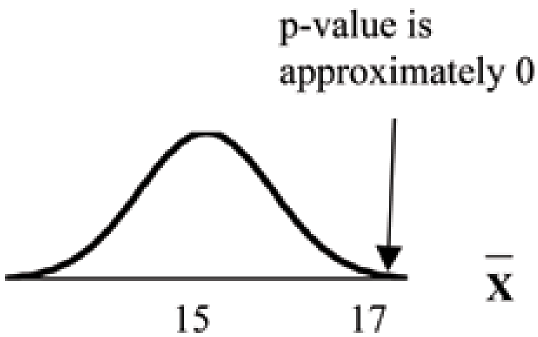
The words **"is more than"** translates as a "$>$ " so "$\mu > 15$" goes into the alternate hypothesis. The null hypothesis must contradict the alternate hypothesis.

Since σ **is known** ($\sigma = 0.5$ cm.), the distribution for the test is normal with mean $\mu = 15$ and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{0.5}{\sqrt{10}} = 0.16$.

Suppose the null hypothesis is true (the average height of the loaves is no more than 15 cm). Then is the average height (17 cm) calculated from the sample unexpectedly large? The hypothesis test works by asking the question how **unlikely** the sample average would be if the null hypothesis were true. The graph shows how far out the sample average is on the normal curve. How far out the sample average is on the normal curve is measured by the p-value. The p-value is the probability that, if we were to take other samples, any other sample average would fall at least as far out as 17 cm.

The p-value, then, is the probability that a sample average is the same or greater than 17 cm. when the population mean is, in fact, 15 cm. We can calculate this probability using the normal distribution for averages from Chapter 7.

[7]This content is available online at <http://http://cnx.org/content/m16995/1.13/>.



p-value $= P\left(\overline{X} > 17\right)$ which is approximately 0.

A p-value of approximately 0 tells us that it is highly unlikely that a loaf of bread rises no more than 15 cm, on the average. That is, almost 0% of all loaves of bread would be at least as high as 17 cm. **purely by CHANCE**. Because the outcome of 17 cm. is so **unlikely (meaning it is happening NOT by chance alone)**, we conclude that the evidence is strongly against the null hypothesis (the average height is at most 15 cm.). There is sufficient evidence that the true average height for the population of the baker's loaves of bread is greater than 15 cm.

9.8 Decision and Conclusion[8]

A systematic way to make a decision of whether to reject or not reject the **null hypothesis** is to compare the **p-value** and a **preset or preconceived α (also called a "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.

When you make a **decision** to reject or not reject H_o, do as follows:

- If $\alpha >$ p-value, reject H_o. The results of the sample data are significant. There is sufficient evidence to conclude that H_o is an incorrect belief and that the **alternative hypothesis**, H_a, may be correct.
- If $\alpha \leq$ p-value, 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_a, may be correct.
- 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 has **failed** to provide sufficient evidence to cast serious doubt about the truthfulness of H_o.

Conclusion: After you make your decision, write a thoughtful **conclusion** about the hypotheses in terms of the given problem.

9.9 Additional Information[9]

- In a **hypothesis test** problem, you may see words such as "the level of significance is 1%." The "1%" is the preconceived or preset α.
- The statistician setting up the hypothesis test selects the value of α to use **before** collecting the sample data.
- **If no level of significance is given, we generally can use $\alpha = 0.05$.**

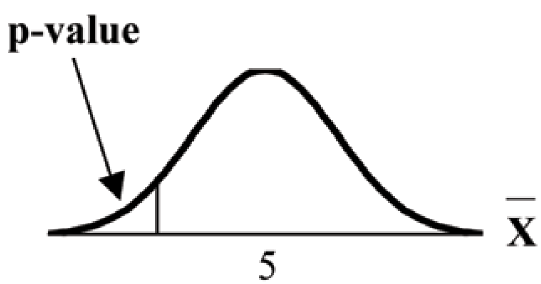
[8]This content is available online at <http://http://cnx.org/content/m16992/1.10/>.
[9]This content is available online at <http://http://cnx.org/content/m16999/1.9/>.

- When you calculate the **p-value** and draw the picture, the p-value is in the left tail, the right tail, or split evenly between the two tails. For this reason, we call the hypothesis test left, right, or two tailed.
- The **alternate hypothesis**, H_a, tells you if the test is left, right, or two-tailed. It is the **key** to conducting the appropriate test.
- H_a **never** has a symbol that contains an equal sign.

The following examples illustrate a left, right, and two-tailed test.

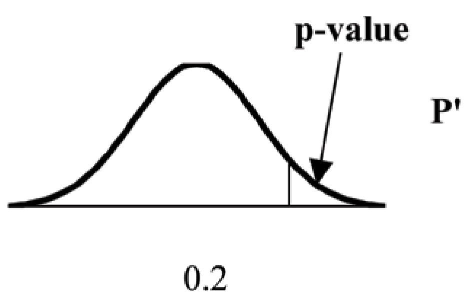
Example 9.8
H_o: $\mu = 5$ H_a: $\mu < 5$

Test of a single population mean. H_a tells you the test is left-tailed. The picture of the p-value is as follows:



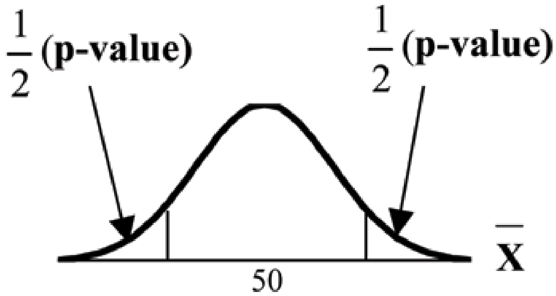
Example 9.9
H_o: $p \leq 0.2$ H_a: $p > 0.2$

This is a test of a single population proportion. H_a tells you the test is **right-tailed**. The picture of the p-value is as follows:



Example 9.10
H_o: $\mu = 50$ H_a: $\mu \neq 50$

This is a test of a single population mean. H_a tells you the test is **two-tailed**. The picture of the p-value is as follows.



$\frac{1}{2}$ (p-value)
$\frac{1}{2}$ (p-value)
50
$\overline{X}$

9.10 Summary of the Hypothesis Test[10]

The **hypothesis test** itself has an established process. This can be summarized as follows:

1. Determine H_o and H_a. Remember, they are contradictory.
2. Determine the random variable.
3. Determine the distribution for the test.
4. Draw a graph, calculate the test statistic, and use the test statistic to calculate the **p-value**. (A z-score and a t-score are examples of test statistics.)
5. Compare the preconceived α with the p-value, make a decision (reject or cannot reject H_o), and write a clear conclusion using English sentences.

Notice that in performing the hypothesis test, you use α and not β. β is needed to help determine the sample size of the data that is used in calculating the p-value. Remember that the quantity $1 - \beta$ is called the **Power of the Test**. A high power is desirable. If the power is too low, statisticians typically increase the sample size while keeping α the same. If the power is low, the null hypothesis might not be rejected when it should be.

9.11 Examples[11]

Example 9.11

Jeffrey, as an eight-year old, **established an average time of 16.43 seconds** for swimming the 25-yard freestyle, with a **standard deviation of 0.8 seconds**. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster by 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 average time was 16 seconds. Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds.** Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume that the swim times for the 25-yard freestyle are normal.

Solution

Set up the Hypothesis Test:

Since the problem is about a mean (average), this is a **test of a single population mean**.

H_o: $\mu = 16.43$ $\qquad$ H_a: $\mu < 16.43$

For Jeffrey to swim faster, his time will be less than 16.43 seconds. The "<" tells you this is left-tailed.

Determine the distribution needed:

Random variable: $\overline{X}$ = the average time to swim the 25-yard freestyle.

Distribution for the test: $\overline{X}$ is normal (population **standard deviation** is known: $\sigma = 0.8$)

$\overline{X} \sim N\left(\mu, \frac{\sigma_X}{\sqrt{n}}\right)$ $\qquad$ Therefore, $\overline{X} \sim N\left(16.43, \frac{0.8}{\sqrt{15}}\right)$

$\mu = 16.43$ comes from H_0 and not the data. $\sigma = 0.8$, and $n = 15$.

Calculate the p-value using the normal distribution for a mean:

[10] This content is available online at <http://http://cnx.org/content/m16993/1.5/>.

[11] This content is available online at <http://http://cnx.org/content/m17005/1.20/>.

p-value $= P\left(\bar{X} < 16\right) = 0.0187$ where the sample mean in the problem is given s 16.

p-value $= 0.0187$ (This is called the **actual level of significance**.) The p-value is the area to the left of the sample mean is given as 16.

Graph:

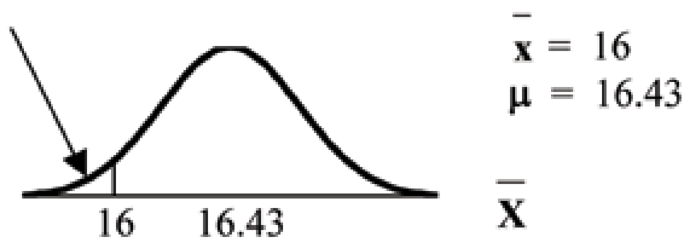


Figure 9.1

$\mu = 16.43$ comes from H_o. Our assumption is $\mu = 16.43$.

Interpretation of the p-value: If H_o **is true**, there is a 0.0187 probability (1.87%) that Jeffrey's mean (or average) time to swim the 25-yard freestyle is 16 seconds or less. Because a 1.87% chance is small, the mean time of 16 seconds or less is not happening randomly. It is a rare event.

Compare α and the p-value:

$\alpha = 0.05$ p-value $= 0.0187$ $\alpha >$ p-value

Make a decision: Since $\alpha >$ p-value, reject H_o.

This means that you reject $\mu = 16.43$. In other words, you do not think Jeffrey swims the 25-yard freestyle in 16.43 seconds but faster with the new goggles.

Conclusion: At the 5% significance level, we conclude that Jeffrey swims faster using the new goggles. The sample data show there is sufficient evidence that Jeffrey's mean time to swim the 25-yard freestyle is less than 16.43 seconds.

The p-value can easily be calculated using the TI-83+ and the TI-84 calculators:

Press `STAT` and arrow over to `TESTS`. Press `1:Z-Test`. Arrow over to `Stats` and press `ENTER`. Arrow down and enter 16.43 for μ_0 (null hypothesis), .8 for σ, 16 for the sample mean, and 15 for n. Arrow down to μ: (alternate hypothesis) and arrow over to $<\mu_0$. Press `ENTER`. Arrow down to `Calculate` and press `ENTER`. The calculator not only calculates the p-value ($p = 0.0187$) but it also calculates the test statistic (z-score) for the sample mean. $\mu < 16.43$ is the alternate hypothesis. Do this set of instructions again except arrow to `Draw` (instead of `Calculate`). Press `ENTER`. A shaded graph appears with $z = -2.08$ (test statistic) and $p = 0.0187$ (p-value). Make sure when you use `Draw` that no other equations are highlighted in $Y =$ and the plots are turned off.

When the calculator does a Z-Test, the `Z-Test` function finds the p-value by doing a normal probability calculation using the **Central Limit Theorem**:

$P\left(\overline{X} < 16 = \texttt{2nd DISTR normcdf}\left(-10\text{^}99, 16, 16.43, 0.8/\sqrt{15}\right)\right.$.

The Type I and Type II errors for this problem are as follows:

The Type I error is to conclude that Jeffrey swims the 25-yard freestyle, on average, in less than 16.43 seconds when, in fact, he actually swims the 25-yard freestyle, on average, in 16.43 seconds. (Reject the null hypothesis when the null hypothesis is true.)

The Type II error is to conclude that Jeffrey swims the 25-yard freestyle, on average, in 16.43 seconds when, in fact, he actually swims the 25-yard freestyle, on average, in less than 16.43 seconds. (Do not reject the null hypothesis when the null hypothesis is false.)

Historical Note: The traditional way to compare the two probabilities, α and the p-value, is to compare their test statistics (z-scores). The calculated test statistic for the p-value is -2.08. (From the Central Limit Theorem, the test statistic formula is $z = \frac{\overline{x}-\mu_X}{\left(\frac{\sigma_X}{\sqrt{n}}\right)}$. For this problem, $\overline{x} = 16$, $\mu_X = 16.43$ from the null hypothesis, $\sigma_X = 0.8$, and $n = 15$.) You can find the test statistic for $\alpha = 0.05$ in the normal table (see **15.Tables** in the Table of Contents). The z-score for an area to the left equal to 0.05 is midway between -1.65 and -1.64 (0.05 is midway between 0.0505 and 0.0495). The z-score is -1.645. Since $-1.645 > -2.08$ (which demonstrates that $\alpha >$ p-value), reject H_o. Traditionally, the decision to reject or not reject was done in this way. Today, comparing the two probabilities α and the p-value is very common and advantageous. For this problem, the p-value, 0.0187 is considerably smaller than α, 0.05. You can be confident about your decision to reject. It is difficult to know that the p-value is traditionally smaller than α by just examining the test statistics. The graph shows α, the p-value, and the two test statistics (z scores).

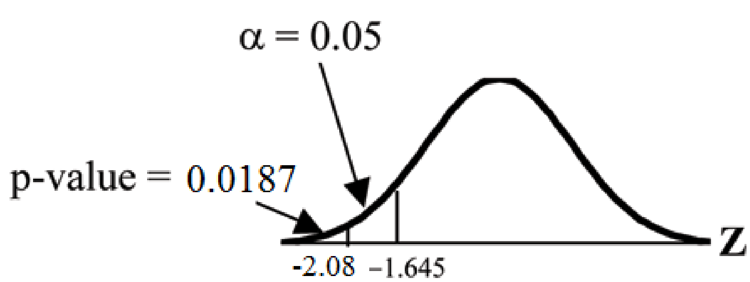


Figure 9.2

Example 9.12
A college football coach thought that his players could bench press an **average of 275 pounds**. It is known that the **standard deviation is 55 pounds**. Three of his players thought that the average was **more than** that amount. They asked **30** of their teammates for their estimated maximum lift on the bench press exercise. The data ranged from 205 pounds to 385 pounds. The actual different

weights were (frequencies are in parentheses) 205(3); 215(3); 225(1); 241(2); 252(2); 265(2); 275(2); 313(2); 316(5); 338(2); 341(1); 345(2); 368(2); 385(1). (Source: data from Reuben Davis, Kraig Evans, and Scott Gunderson.)

Conduct a hypothesis test using a 2.5% level of significance to determine if the bench press average is **more than 275 pounds**.

Solution
Set up the Hypothesis Test:

Since the problem is about a mean (average), this is a **test of a single population mean**.

H_o: $\mu = 275$ H_a: $\mu > 275$ This is a right-tailed test.

Calculating the distribution needed:

Random variable: $\overline{X}$ = the average weight lifted by the football players.

Distribution for the test: It is normal because σ is known.

$\overline{X} \sim N\left(275, \frac{55}{\sqrt{30}}\right)$

$\overline{x} = 286.2$ pounds (from the data).

$\sigma = 55$ pounds **(Always use σ if you know it.)** We assume $\mu = 275$ pounds unless our data shows us otherwise.

Calculate the p-value using the normal distribution for a mean:

p-value $= P(\ \overline{X} > 286.2 = 0.1323$ where the sample mean is calculated as 286.2 pounds from the data.

Interpretation of the p-value: If H_o is true, then there is a 0.1323 probability (13.23%) that the football players can lift a mean (or average) weight of 286.2 pounds or more. Because a 13.23% chance is large enough, a mean weight lift of 286.2 pounds or more is happening randomly and is not a rare event.

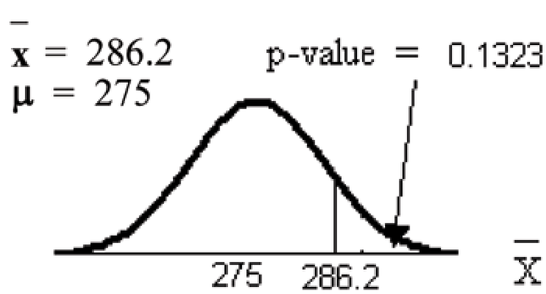


Figure 9.3

Compare α and the p-value:

$\alpha = 0.025$ p-value $= 0.1323$

Make a decision: Since α<p-value, do not reject H_o.

Conclusion: At the 2.5% level of significance, from the sample data, there is not sufficient evidence to conclude that the true mean weight lifted is more than 275 pounds.

The p-value can easily be calculated using the TI-83+ and the TI-84 calculators:

Put the data and frequencies into lists. Press `STAT` and arrow over to `TESTS`. Press `1:Z-Test`. Arrow over to `Data` and press `ENTER`. Arrow down and enter 275 for μ_0, 55 for σ, the name of the list where you put the data, and the name of the list where you put the frequencies. Arrow down to μ : and arrow over to $> \mu_0$. Press `ENTER`. Arrow down to `Calculate` and press `ENTER`. The calculator not only calculates the p-value ($p = 0.1331$, a little different from the above calculation - in it we used the sample mean rounded to one decimal place instead of the data) but it also calculates the test statistic (z-score) for the sample mean, the sample mean, and the sample standard deviation. $\mu > 275$ is the alternate hypothesis. Do this set of instructions again except arrow to `Draw` (instead of `Calculate`). Press `ENTER`. A shaded graph appears with $z = 1.112$ (test statistic) and $p = 0.1331$ (p-value). Make sure when you use `Draw` that no other equations are highlighted in $Y =$ and the plots are turned off.

Example 9.13

Statistics students believe that the average score on the first statistics test is 65. A statistics instructor thinks the average score is higher than 65. He samples ten statistics students and obtains the scores 65; 65; 70; 67; 66; 63; 63; 68; 72; 71. He performs a hypothesis test using a 5% level of significance. The data are from a normal distribution.

Solution

Set up the Hypothesis Test:

A 5% level of significance means that $\alpha = 0.05$. This is a test of a **single population mean**.

H_o: $\mu = 65$ $\qquad$ H_a: $\mu > 65$

Since the instructor thinks the average score is higher, use a "$>$ ". The "$>$ " means the test is right-tailed.

Determine the distribution needed:

Random variable: $\overline{X}$ = average score on the first statistics test.

Distribution for the test: If you read the problem carefully, you will notice that there is **no population standard deviation given**. You are only given $n = 10$ sample data values. Notice also that the data come from a normal distribution. This means that the distribution for the test is a student-t.

Use t_{df}. Therefore, the distribution for the test is t_9 where $n = 10$ and $\text{df} = 10 - 1 = 9$.

Calculate the p-value using the Student-t distribution:

p-value $= P(\ \overline{X} > 67 = 0.0396$ where the sample mean and sample standard deviation are calculated as 67 and 3.1972 from the data.

Interpretation of the p-value: If the null hypothesis is true, then there is a 0.0396 probability (3.96%) that the sample mean is 67 or more.

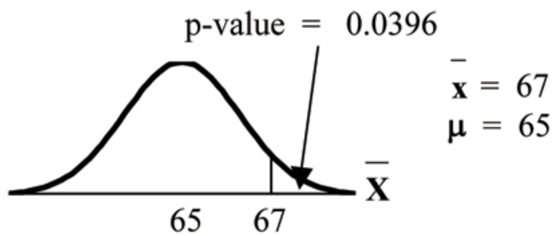


Figure 9.4

Compare α and the p-value:

Since $\alpha = .05$ and p-value $= 0.0396$. Therefore, $\alpha >$ p-value.

Make a decision: Since $\alpha >$ p-value, reject H_o.

This means you reject $\mu = 65$. In other words, you believe the average test score is more than 65.

Conclusion: At a 5% level of significance, the sample data show sufficient evidence that the mean (average) test score is more than 65, just as the math instructor thinks.

The p-value can easily be calculated using the TI-83+ and the TI-84 calculators:

Put the data into a list. Press `STAT` and arrow over to `TESTS`. Press `2:T-Test`. Arrow over to `Data` and press `ENTER`. Arrow down and enter 65 for μ_0, the name of the list where you put the data, and 1 for `Freq:`. Arrow down to μ : and arrow over to $> \mu_0$. Press `ENTER`. Arrow down to `Calculate` and press `ENTER`. The calculator not only calculates the p-value ($p = 0.0396$) but it also calculates the test statistic (t-score) for the sample mean, the sample mean, and the sample standard deviation. $\mu > 65$ is the alternate hypothesis. Do this set of instructions again except arrow to `Draw` (instead of `Calculate`). Press `ENTER`. A shaded graph appears with $t = 1.9781$ (test statistic) and $p = 0.0396$ (p-value). Make sure when you use `Draw` that no other equations are highlighted in $Y =$ and the plots are turned off.

Example 9.14
Joon believes that 50% of first-time brides in the United States are younger than their grooms. She performs a hypothesis test to determine if the percentage is **the same or different from 50%**. Joon samples **100 first-time brides** and **53** reply that they are younger than their grooms. For the hypothesis test, she uses a 1% level of significance.

Solution
Set up the Hypothesis Test:

The 1% level of significance means that $\alpha = 0.01$. This is a **test of a single population proportion**.

H_o: $p = 0.50$ $\qquad$ H_a: $p \neq 0.50$

The words **"is the same or different from"** tell you this is a two-tailed test.

Calculate the distribution needed:

Random variable: P' = the percent of of first-time brides who are younger than their grooms.

Distribution for the test: The problem contains no mention of an average. The information is given in terms of percentages. Use the distribution for P', the estimated proportion.

$P' \sim N\left(p, \sqrt{\frac{p \cdot q}{n}}\right)$ Therefore, $P' \sim N\left(0.5, \sqrt{\frac{0.5 \cdot 0.5}{100}}\right)$ where $p = 0.50$, $q = 1 - p = 0.50$, and $n = 100$.

Calculate the p-value using the normal distribution for proportions:

p-value $= P\,(P' < 0.47$ or $P' > 0.53 = 0.5485$

where $x = 53$, $p' = \frac{x}{n} = \frac{53}{100} = 0.53$.

Interpretation of the p-value: If the null hypothesis is true, there is 0.5485 probability (54.85%) that the sample (estimated) proportion p' is 0.53 or more OR 0.47 or less (see the graph below).

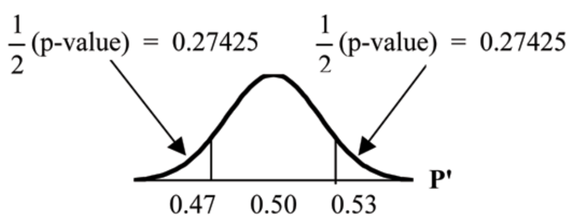


Figure 9.5

$\mu = p = 0.50$ comes from H_o, the null hypothesis.

p'= 0.53. Since the curve is symmetrical and the test is two-tailed, the p' for the left tail is equal to $0.50 - 0.03 = 0.47$ where $\mu = p = 0.50$. (0.03 is the difference between 0.53 and 0.50.)

Compare α and the p-value:

Since $\alpha = 0.01$ and p-value $= 0.5485$. Therefore, $\alpha <$ p-value.

Make a decision: Since $\alpha <$p-value, you cannot reject H_o.

Conclusion: At the 1% level of significance, the sample data do not show sufficient evidence that the percentage of first-time brides who are younger than their grooms is different from 50%.

The p-value can easily be calculated using the TI-83+ and the TI-84 calculators:

Press `STAT` and arrow over to `TESTS`. Press `5:1-PropZTest`. Enter .5 for p_0 and 100 for n. Arrow down to `Prop` and arrow to `not equals` p_0. Press `ENTER`. Arrow down to `Calculate` and press `ENTER`. The calculator calculates the p-value ($p = 0.5485$) and the test statistic (z-score). `Prop not equals` .5 is the alternate hypothesis. Do this set of instructions again except arrow to `Draw` (instead of `Calculate`). Press `ENTER`. A shaded graph appears with $z = 0.6$ (test statistic) and $p = 0.5485$

(p-value). Make sure when you use `Draw` that no other equations are highlighted in $Y =$ and the plots are turned off.

The Type I and Type II errors are as follows:

The Type I error is to conclude that the proportion of first-time brides that are younger than their grooms is different from 50% when, in fact, the proportion is actually 50%. (Reject the null hypothesis when the null hypothesis is true).

The Type II error is to conclude that the proportion of first-time brides that are younger than their grooms is equal to 50% when, in fact, the proportion is different from 50%. (Do not reject the null hypothesis when the null hypothesis is false.)

Example 9.15
Problem 1
Suppose a consumer group suspects that the proportion of households that have three cell phones is not known to be 30%. A cell phone company has reason to believe that the proportion is 30%. Before they start a big advertising campaign, they conduct a hypothesis test. Their marketing people survey 150 households with the result that 43 of the households have three cell phones.

Solution
Set up the Hypothesis Test:

H_o: $p = 0.30$ $\qquad$ H_a: $p \neq 0.30$

Determine the distribution needed:

The **random variable** is P' = proportion of households that have three cell phones.

The **distribution** for the hypothesis test is $P' \sim N\left(0.30, \sqrt{\frac{0.30 \cdot 0.70}{150}}\right)$

Problem 2
The value that helps determine the p-value is p'. Calculate p'.

Problem 3
What is a **success** for this problem?

Problem 4
What is the level of significance?

Draw the graph for this problem. Draw the horizontal axis. Label and shade appropriately.

Problem 5
Calculate the p-value.

Problem 6
Make a decision. _____________(Reject/Do not reject) H_0 because____________.

The next example is a poem written by a statistics student named Nicole Hart. The solution to the problem follows the poem. Notice that the hypothesis test is for a single population proportion. This means that the null and alternate hypotheses use the parameter p. The distribution for the test is normal. The estimated proportion p' is the proportion of fleas killed to the total fleas found on Fido. This is sample information. The problem gives a preconceived $\alpha = 0.01$, for comparison, and a 95% confidence interval computation. The poem is clever and humorous, so please enjoy it!

NOTE: Hypothesis testing problems consist of multiple steps. To help you do the problems, solution sheets are provided for your use. Look in the Table of Contents Appendix for the topic "Solution Sheets." If you like, use copies of the appropriate solution sheet for homework problems.

Example 9.16

```
    My dog has so many fleas,
They do not come off with ease.
As for shampoo, I have tried many types
Even one called Bubble Hype,
Which only killed 25% of the fleas,
Unfortunately I was not pleased.

I've used all kinds of soap,
Until I had give up hope
Until one day I saw
An ad that put me in awe.

A shampoo used for dogs
Called GOOD ENOUGH to Clean a Hog
Guaranteed to kill more fleas.

I gave Fido a bath
And after doing the math
His number of fleas
Started dropping by 3's!

Before his shampoo
I counted 42.
At the end of his bath,
I redid the math
And the new shampoo had killed 17 fleas.
So now I was pleased.

Now it is time for you to have some fun
With the level of significance being .01,
You must help me figure out
Use the new shampoo or go without?
```

Solution
Set up the Hypothesis Test:

H_o: $p = 0.25$ H_a: $p > 0.25$

Determine the distribution needed:

In words, CLEARLY state what your random variable $\overline{X}$ or P' represents.

P' = The proportion of fleas that are killed by the new shampoo

State the distribution to use for the test.

Normal: $N\left(0.25, \sqrt{\frac{(0.25)(1-0.25)}{42}}\right)$

Test Statistic: $z = 2.3163$

Calculate the p-value using the normal distribution for proportions:

p-value $=0.0103$

In 1 – 2 complete sentences, explain what the p-value means for this problem.

If the null hypothesis is true (the proportion is 0.25), then there is a 0.0103 probability that the sample (estimated) proportion is 0.4048 $\left(\frac{17}{42}\right)$ or more.

Use the previous information to sketch a picture of this situation. CLEARLY, label and scale the horizontal axis and shade the region(s) corresponding to the p-value.

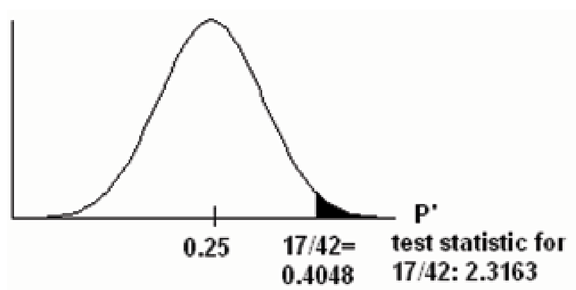


Figure 9.6

Compare α and the p-value:

Indicate the correct decision ("reject" or "do not reject" the null hypothesis), the reason for it, and write an appropriate conclusion, using COMPLETE SENTENCES.

alpha	decision	reason for decision
0.01	Do not reject H_o	$\alpha <$p-value

Table 9.3

Conclusion: At the 1% level of significance, the sample data do not show sufficient evidence that the percentage of fleas that are killed by the new shampoo is more than 25%.

Construct a 95% Confidence Interval for the true mean or proportion. Include a sketch of the graph of the situation. Label the point estimate and the lower and upper bounds of the Confidence Interval.

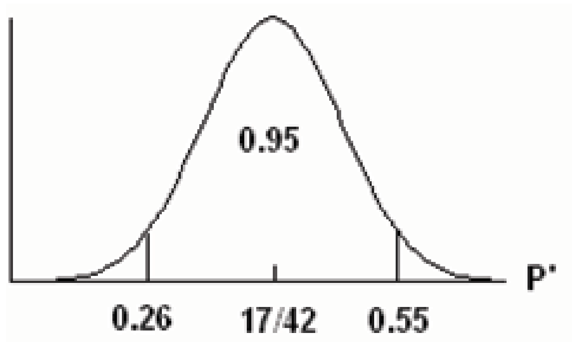


Figure 9.7

Confidence Interval: $(0.26, 0.55)$ We are 95% confident that the true population proportion p of fleas that are killed by the new shampoo is between 26% and 55%.

NOTE: This test result is not very definitive since the p-value is very close to alpha. In reality, one would probably do more tests by giving the dog another bath after the fleas have had a chance to return.

9.12 Summary of Formulas[12]

H_o and H_a are contradictory.

If H_o has:	equal $(=)$	greater than or equal to $(\geq)$	less than or equal to $(\leq)$
then H_a has:	not equal $(\neq)$ **or** greater than $(>)$ **or** less than $(<)$	less than $(<)$	greater than $(>)$

Table 9.4

If $\alpha \leq$ p-value, then do not reject H_o.

If $\alpha >$ p-value, then reject H_o .

α is preconceived. Its value is set before the hypothesis test starts. The p-value is calculated from the data.

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

β = probability of a Type II error = P(Type II error) = probability of not rejecting the null hypothesis when the null hypothesis is false.

If there is no given preconceived α, then use $\alpha = 0.05$.

Types of Hypothesis Tests

- Single population mean, **known** population variance (or standard deviation): **Normal test**.
- Single population mean, **unknown** population variance (or standard deviation): **Student-t test**.
- Single population proportion: **Normal test**.

[12]This content is available online at <http://http://cnx.org/content/m16996/1.7/>.