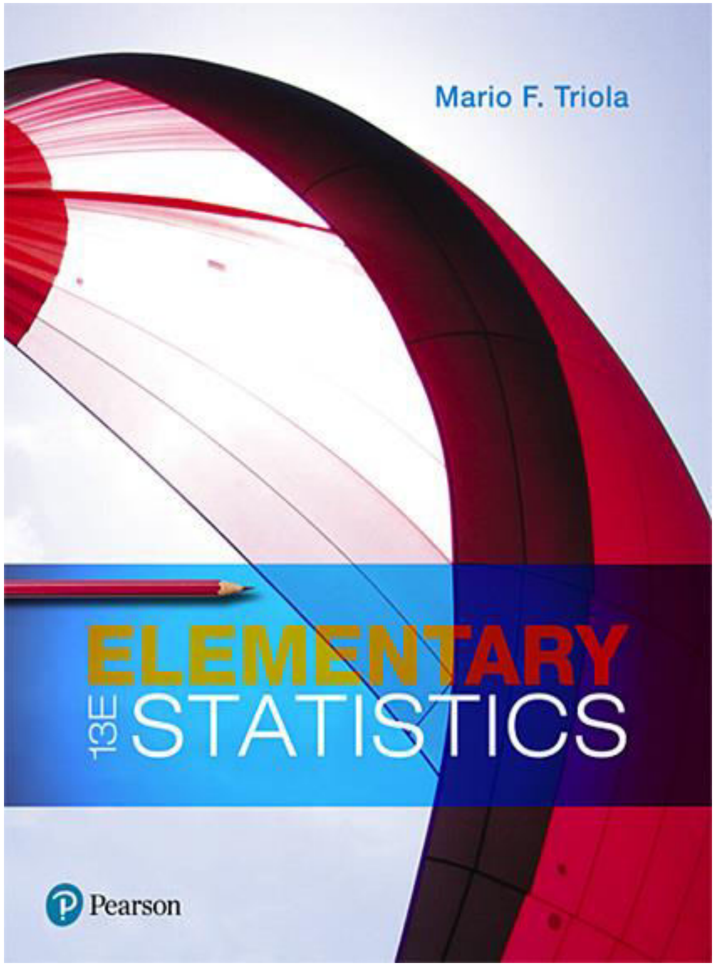
Elementary Statistics

Thirteenth Edition



Chapter 8

Hypothesis Testing

Hypothesis Testing

8-1 Basics of Hypothesis Testing

8-2 Testing a Claim about a Proportion

8-3 Testing a Claim About a Mean

8-4 Testing a Claim About a Standard Deviation or Variance

Key Concept

In this section we present key components of a formal hypothesis test. The concepts in this section are general and apply to hypothesis tests involving proportions, means, or standard deviations or variances.

Hypothesis and Hypothesis Test

- Hypothesis
  - In statistics, a **hypothesis** is a claim or statement about a property of a population.
- Hypothesis Test
  - A **hypothesis test** (or **test of significance**) is a procedure for testing a claim about a property of a population.

Example: Majority of Consumers are not Comfortable with Drone Deliveries (1 of 6)

1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% (or 545) of them responded with “no.” Using p to denote the proportion of consumers not comfortable with drone deliveries, the “majority” claim is equivalent to the claim that the proportion is greater than half, or $p > 0.5$. The expression $p > 0.5$ is the symbolic form of the original claim.

Example: Majority of Consumers are not Comfortable with Drone Deliveries (2 of 6)

The Big Picture We have the claim that the population proportion p is such that $p > 0.5$. Among 1009 consumers, how many do we need to get a **significantly high** number who are not comfortable with drone delivery?

- A result of 506 (or 50.1%) is just barely more than half, so 506 is clearly **not significantly high.**
- A result of 1006 (or 99.7%) is clearly **significantly high.** But what about the result of 545 (or 54.0%) that was actually obtained in the Pitney Bowes survey?
- Is 545 (or 54.0%) **significantly high?** The method of hypothesis testing allows us to answer that key question.

Example: Majority of Consumers are not Comfortable with Drone Deliveries (3 of 6)

Using Technology It is easy to obtain hypothesis-testing results using technology. The accompanying screen displays show results from four different technologies, **so we can use computers or calculators to do all of the computational heavy lifting.**

Example: Majority of Consumers are not Comfortable with Drone Deliveries (4 of 6)

Using Technology

Statdisk

Alternative Hypothesis:
p > p(hyp)

Sample proportion: 0.5401388
Test Statistic, z: 2.5500
Critical z: 1.6449
P-Value: 0.0054

90% Confidence interval:
0.5143312 < p < 0.5659463

Minitab

```
Test of p = 0.5 vs p > 0.5

Sample     X     N  Sample p  95% Lower Bound  Z-Value  P-Value
1        545  1009  0.540139         0.514331     2.55    0.005

Using the normal approximation.
```

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
        1-PropZTest
prop>.5
z=2.549995628
p=.005386238
p̂=.5401387512
n=1009
```

StatCrunch

Hypothesis test results:
p : Proportion of successes
$H_0 : p = 0.5$
$H_A : p > 0.5$

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	545	1009	0.54013875	0.015740714	2.5499956	0.0054

Example: Majority of Consumers are not Comfortable with Drone Deliveries (5 of 6)

Using Technology

Examining the four screen displays, we see some common elements. They all display a "test statistic" of $z = 2.55$ (rounded), and they all include a "*P*-value" of 0.005 (rounded).

Focus on **understanding** how the hypothesis-testing procedure works and learn the associated terminology. Only then will results from technology make sense.

Example: Majority of Consumers are not Comfortable with Drone Deliveries (6 of 6)

Significance

Hypothesis tests are also called **tests of significance**. In Section 4-1 we used probabilities to determine when sample results are **significantly low** or **significantly high**. This chapter formalizes those concepts in a unified procedure that is used often throughout many different fields of application.

Null Hypothesis

- Null Hypothesis
  - The **null hypothesis** (denoted by H_0) is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is **equal to** some claimed value.

Alternative Hypothesis

- Alternative Hypothesis
  - The **alternative hypothesis** (denoted by H_1 or H_a or H_A) is a statement that the parameter has a value that somehow differs from the null hypothesis. For the methods of this chapter, the symbolic form of the alternative hypothesis must use one of these symbols: $<$, $>$, $\neq$.

Procedure for Hypothesis Tests

1. Identify the Claim
Identify the claim to be tested and express it in symbolic form.

↓

2. Give Symbolic Form
Give the symbolic form that must be true when the original claim is false.

↓

3. Identify Null and Alternative Hypothesis
Consider the two symbolic expressions obtained so far:

- **Alternative hypothesis H_1** is the one *NOT* containing equality, so H_1 uses the symbol $>$ or $<$ or $\neq$.
- **Null hypothesis H_0** is the symbolic expression that the parameter equals the fixed value being considered.

↓

4. Select Significance Level
Select the **significance level** α based on the seriousness of a type I error. Make α small if the consequences of rejecting a true H_0 are severe.

- The values of 0.05 and 0.01 are very common.

↓

5. Identify the Test Statistic
Identify the test statistic that is relevant to the test and determine its sampling distribution (such as normal, *t*, chi-square).

↓

***P*-Value Method**

6. Find Values
Find the value of the **test statistic** and the ***P*-value** (see Figure 8-3). Draw a graph and show the test statistic and *P*-value.

↓

7. Make a Decision

- **Reject H_0** if *P*-value $\leq \alpha$.
- **Fail to reject H_0** if *P*-value $> \alpha$.

↓

Critical Value Method

6. Find Values
Find the value of the **test statistic** and the **critical values**. Draw a graph showing the test statistic, critical value(s) and critical region.

↓

7. Make a Decision

- **Reject H_0** if the test statistic is in the critical region.
- **Fail to reject H_0** if the test statistic is not in the critical region.

↓

8. Restate Decision in Nontechnical Terms
Restate this previous decision in simple nontechnical terms, and address the original claim.

Confidence Interval Method

Confidence Interval Method

Construct a confidence interval with a confidence level selected as in Table 8-1.

Table 8-1 Confidence Level for Confidence Interval

Significance Level for Hypothesis Test	Two-Tailed Test	One-Tailed Test
0.01	99%	98%
0.05	95%	90%
0.10	90%	80%

Because a confidence interval estimate of a population parameter contains the likely values of that parameter, reject a claim that the population parameter has a value that is not included in the confidence interval.

Use the Original Claim to Create a Null Hypothesis H_0 and an Alternative Hypothesis H_1

Step 1. Identify the claim to be tested and express it in symbolic form.

Step 2. Give the symbolic form that must be true when the original claim is false.

Step 3. Consider the two symbolic expressions obtained so far:

- **Alternative hypothesis H_1** is the one **NOT** containing equality, so H_1 uses the symbol $<$ or $>$ or $\neq$.
- **Null hypothesis H_0** is the symbolic expression that the parameter **equals** the fixed value being considered.

Example: Drone Delivery (1 of 8)

Given the claim that "the majority of consumers are uncomfortable with drone delivery," we can apply Steps 1, 2, and 3 as follows.

Example: Drone Delivery (2 of 8)

Step 1: Identify the claim to be tested and express it in symbolic form. Using p to denote the probability of selecting a consumer uncomfortable with drone delivery, the claim that "the majority is uncomfortable with drone delivery" can be expressed in symbolic form as $p > 0.5$.

Step 2: Give the symbolic form that must be true when the original claim is false. If the original claim of $p > 0.5$ is false, then $p \le 0.5$ must be true.

Example: Drone Delivery (3 of 8)

Step 3: This step is in two parts: Identify the alternative hypothesis H_1 and identify the null hypothesis H_0.

- Identify H_1: Using the two symbolic expressions $p > 0.5$ and $p \le 0.5$, the alternative hypothesis H_1 is the one that does not contain equality. Of those two expressions, $p > 0.5$ does not contain equality, so we get H_1: $p > 0.5$
- Identify H_0: The null hypothesis H_0 is the symbolic expression that the parameter **equals** the fixed value being considered, so we get H_0: $p = 0.5$

Step 4: Significance Level α

- Significance Level
  - The **significance level** α for a hypothesis test is the probability value used as the cutoff for determining when the sample evidence constitutes **significant** evidence against the null hypothesis. By its nature, the significance level α is the probability of mistakenly rejecting the null hypothesis when it is true:

Significance level $\alpha = P$ (rejecting H_0 when H_0 is true)

Select the Significance Level α

Step 4. The **significance level α** is the same α introduced in sections 7-1, where we defined “critical value”. Common choices for α are 0.05, 0.01, and 0.10; 0.05 is most common.

Identify the test statistic that is relevant to the test and determine its sampling distribution (such as normal, *t*, χ^2) (1 of 2)

Step 5. Identify the test statistic that is relevant to the test and determine its sampling distribution (such as normal, t, χ^2). The table on the following slide lists parameters along with the corresponding sampling distributions.

Identify the test statistic that is relevant to the test and determine its sampling distribution (such as normal, *t*, χ^2) (2 of 2)

Parameter	Sampling Distribution	Requirements	Test Statistic
Proportion p	Normal (z)	$np \geq 5$ and $nq \geq 5$	$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$
Mean $\boldsymbol{\mu}$	t	$\boldsymbol{\sigma}$ not known and normally distributed population or $\boldsymbol{\sigma}$ not known and $n > 30$	$t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}}$
Mean $\boldsymbol{\mu}$	Normal (z)	$\boldsymbol{\sigma}$ known and normally distributed population or $\boldsymbol{\sigma}$ known and $n > 30$	$z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$
St. dev. $\boldsymbol{\sigma}$ or variance $\boldsymbol{\sigma}^2$	χ^2	Strict requirement: normally distributed population	$\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

Example: Drone Delivery (4 of 8)

The claim $p > 0.5$ is a claim about the population proportion p, so use the normal distribution, provided that the requirements are satisfied. (With $n = 1009$, $p = 0.5$, and $q = 0.5$ from the ongoing example, $np \geq 5$ and $nq \geq 5$ are both true.)

Find the Value of the Test Statistic, Then Find Either the *P*-Value or the Critical Values(s)

Step 6. Find the value of the **test statistic** and the ***P*-value** or **critical value(s)**.

Test Statistic

- Test Statistic
  - The **test statistic** is a value used in making a decision about the null hypothesis. It is found by converting the sample statistic (such as $\hat{p}$, $\bar{x}$, or s) to a score (such as z, t, or χ^2) with the assumption that the null hypothesis is true.

Example: Drone Delivery (5 of 8)

We have a claim made about the population proportion p, we have n = 1009 and x = 545, so $\hat{p} = \frac{x}{n} = 0.540.$

With the null hypothesis of H_0: p = 0.5, we are working with the assumption that p = 0.5, and it follows that $q = 1 - p = 0.5$. We can evaluate the test statistic as shown below. The test statistic of z = 2.55 from each of the previous technology displays is more accurate than the result of z = 2.54 shown below.

(If we replace 0.540 with $\frac{545}{1009} = 0.54013875,$ we get z = 2.55.)

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.540 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1009}}} = 2.54$$

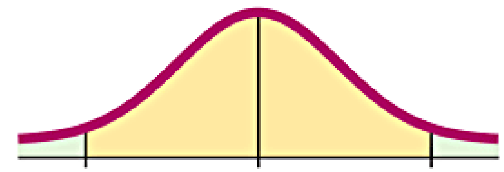
Pearson

Critical Region

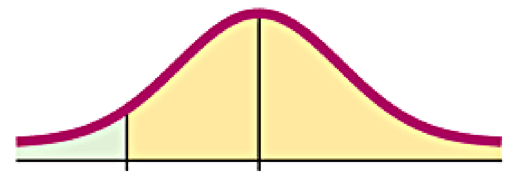
- Critical Region
  - The **critical region** (or **rejection region**) is the area corresponding to all values of the test statistic that cause us to reject the null hypothesis.

Two-Tailed, Left-Tailed, Right-Tailed

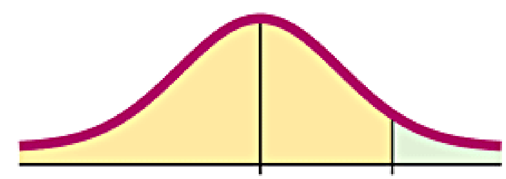
- **Two-tailed test:** The critical region is in the two extreme regions (tails) under the curve.
- **Left-tailed test:** The critical region is in the extreme left region (tail) under the curve.
- **Right-tailed test:** The critical region is in the extreme right region (tail) under the curve.



Sign used in H_1: $\neq$
Two-tailed test



Sign used in H_1: $<$
Left-tailed test



Sign used in H_1: $>$
Right-tailed test

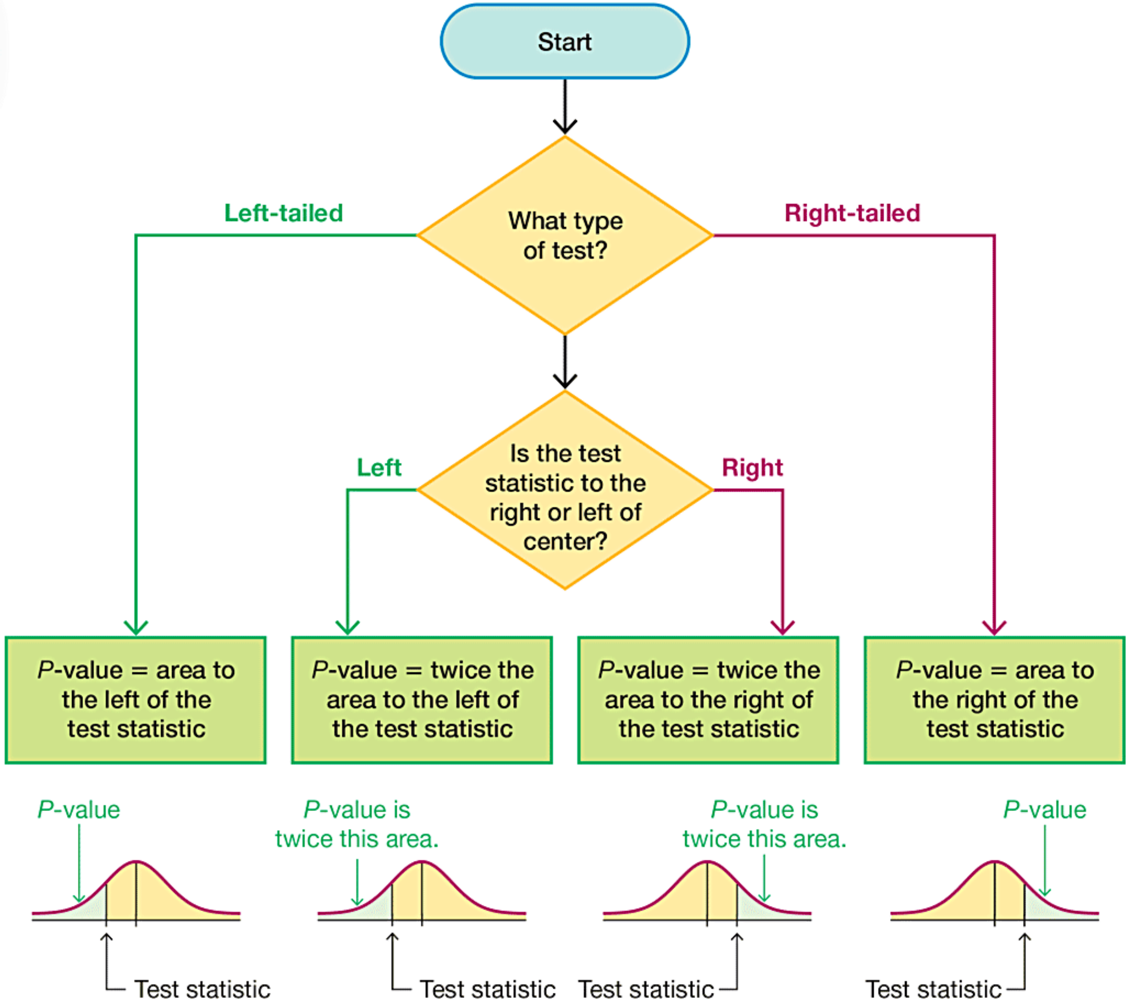
P-Value Method

- *P*-Value Method
  - In a hypothesis test, the ***P*-value** is the probability of getting a value of the test statistic that is **at least as extreme** as the test statistic obtained from the sample data, assuming that the null hypothesis is true.

Example: Drone Delivery (6 of 8)

Using the data from the previous problem, the test statistic is $z = 2.55$, and it has a normal distribution area of 0.0054 to its right, so a right-tailed test with test statistic $z = 2.55$ has a P-value of 0.0054. See the different technology displays given earlier, and note that each of them provides the same P-value of 0.005 after rounding.

Finding *P*-Values



Caution

Don’t confuse a *P*-value with the parameter *p* or the statistic $\hat{p}$. Know the following notation:

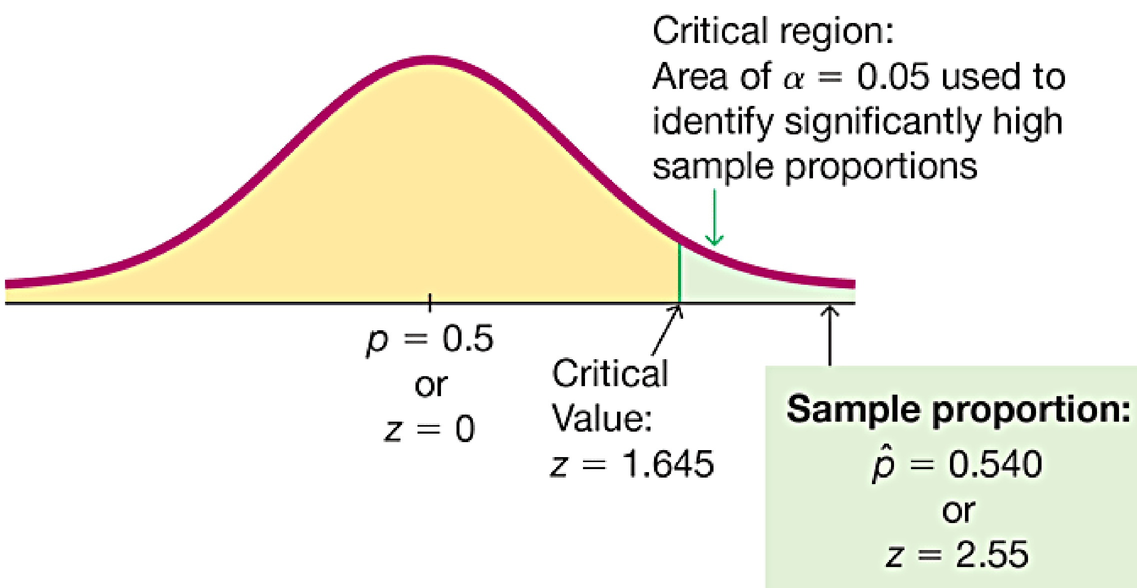
- ***P*-value** = probability of a test statistic at least as extreme as the one obtained
- ***p*** = population proportion
- $\hat{p}$ = sample proportion

Critical Value Method

- Critical Values
  - In a hypothesis test, the **critical value(s)** separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis.

Example: Drone Delivery (7 of 8)

The critical region is shaded in green. The figure shows that with a significance level of $\alpha = 0.05$, the critical value is $z = 1.645$.



Pearson

Make a decision to Either Reject H_0 or Fail to Reject H_0

Step 7. Make a decision to either reject H_0 or fail to reject H_0.

Decision Criteria for the *P*-Value Method:

- If *P*-value $\leq \alpha$, reject H_0 ("If the *P* is low, the null must go.")
- If *P*-value $> \alpha$, fail to reject H_0.

Restate the Decision Using Simple and Nontechnical Terms

Step 8. Restate the decision using simple and nontechnical terms.

Without using technical terms not understood by most people, state a final conclusion that addresses the original claim with wording that can be understood by those without knowledge of statistical procedures.

Example: Drone Delivery (8 of 8)

There is sufficient evidence to support the claim that the majority of consumers are uncomfortable with drone deliveries.

Restate the Decision Using Simple and Nontechnical Terms (1 of 4)

Wording the Final Conclusion

For help in wording the final conclusion, refer to the table on the next slide, which lists the four possible circumstances and their corresponding conclusions.

Note that only the first case leads to wording indicating **support** for the original conclusion. If you want to support some claim, state it in such a way that it becomes the alternative hypothesis, and then hope that the null hypothesis gets rejected.

Restate the Decision Using Simple and Nontechnical Terms (2 of 4)

Wording the Final Conclusion

Condition	Conclusion
Original claim does not include equality, and you reject H_0.	"There is sufficient evidence to **support** the claim that … (original claim)"
Original claim does not include equality, and you fail to reject H_0.	"There is not sufficient evidence to **support** the claim that … (original claim)"
Original claim includes equality, and you reject H_0.	"There is sufficient evidence to warrant **rejection** of the claim that … (original claim)."
Original claim includes equality, and you fail to reject H_0.	"There is not sufficient evidence to warrant **rejection** of the claim that … (original claim)."

Restate the Decision Using Simple and Nontechnical Terms (3 of 4)

Accept or Fail to Reject?

We should say that we “fail to reject the null hypothesis” instead of saying that we “accept the null hypothesis.” The term **accept** is misleading, because it implies incorrectly that the null hypothesis has been proved, but we can never prove a null hypothesis. The phrase **fail to reject** says more correctly that the available evidence isn’t strong enough to warrant rejection of the null hypothesis.

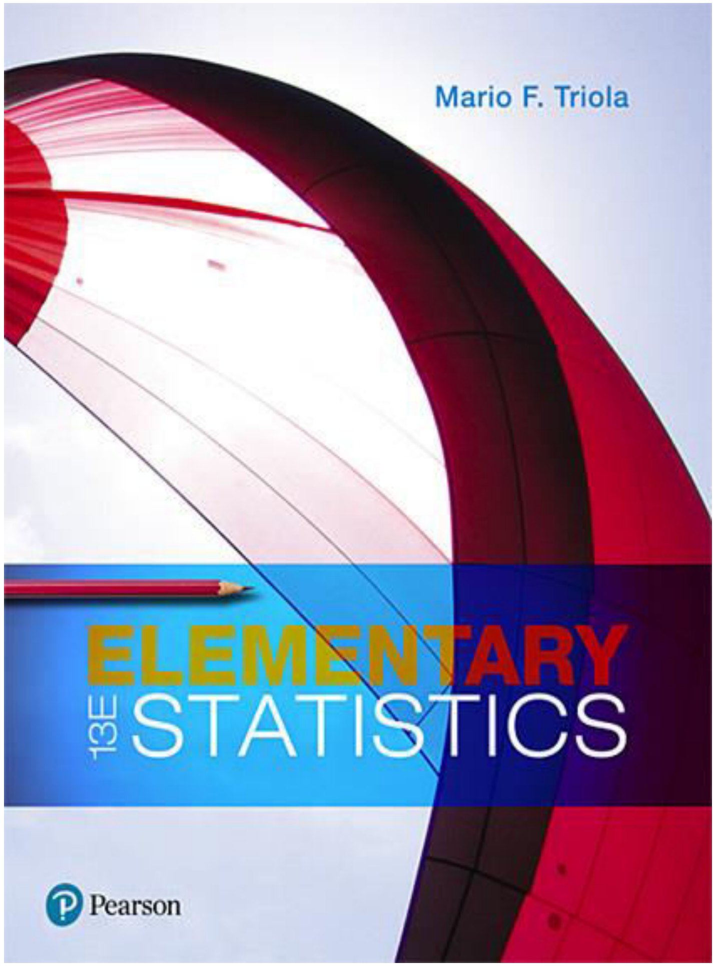
Restate the Decision Using Simple and Nontechnical Terms (4 of 4)

Multiple Negatives

Final conclusions can include as many as three negative terms. For such confusing conclusions, it is better to restate them to be understandable. Instead of saying that “there is not sufficient evidence to warrant rejection of the claim of no difference between 0.5 and the population proportion,” a better statement would be this: “Until stronger evidence is obtained, continue to assume that the population proportion is equal to 0.5.”

Elementary Statistics

Thirteenth Edition



Chapter 8

Hypothesis Testing

Hypothesis Testing

8-1 Basics of Hypothesis Testing

8-2 Testing a Claim about a Proportion

8-3 Testing a Claim About a Mean

8-4 Testing a Claim About a Standard Deviation or Variance

Key Concept

This section presents methods for conducting a formal hypothesis test of a claim made about a population standard deviation σ or population variance σ^2. The methods of this section use the chi-square distribution.

Testing Claims about σ or σ^2: Objective

Objective

Conduct a hypothesis test of a claim made about a population standard deviation σ or population variance σ^2.

Testing Claims about σ or σ^2: Notation

Notation

n = sample size

s = **sample** standard deviation

σ = **population** standard deviation

s^2 = **sample** variance

σ^2 = **population** variance

Testing Claims about σ or σ^2: Requirements

Requirements

1. The sample is a simple random sample.
2. The population has a normal distribution. (This is a fairly strict requirement.)

Testing Claims about σ or σ^2: Test Statistic

Test Statistic

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \text{ (round to three decimal places)}$$

***P*-values:** Use technology or Table A-4 with degrees of freedom: df = $n - 1$.

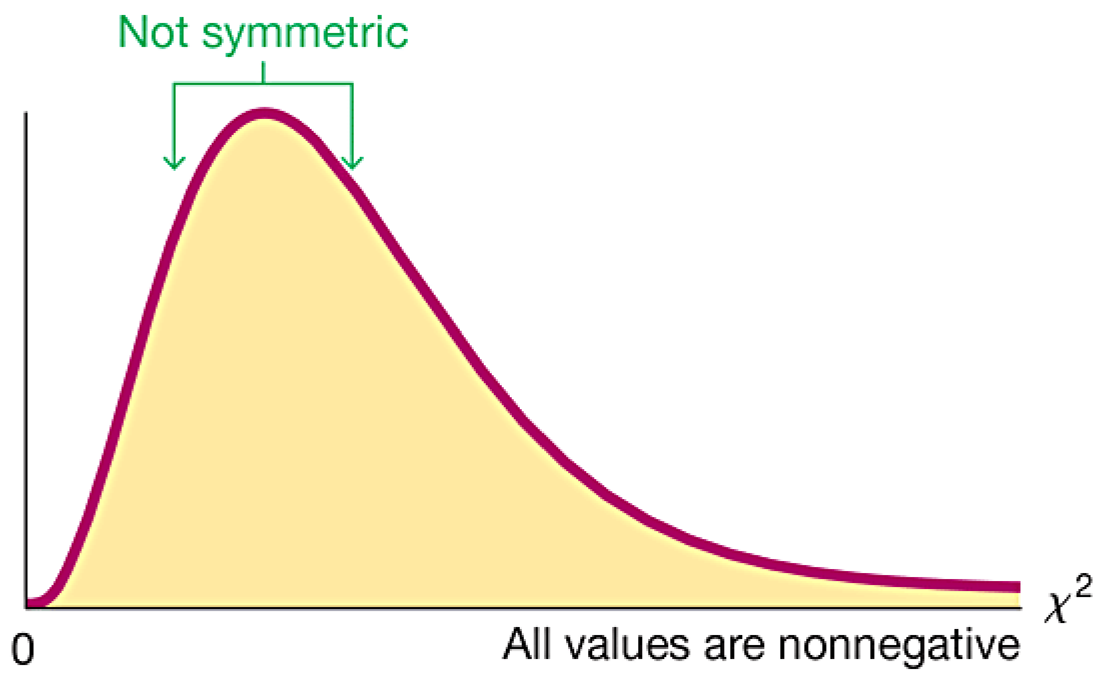
Critical values: Use Table A-4 with degrees of freedom df = $n - 1$.

Equivalent Methods

When testing claims about σ or σ^2, the *P*-value method, the critical value method, and the confidence interval method are all equivalent in the sense that they will always lead to the same conclusion.

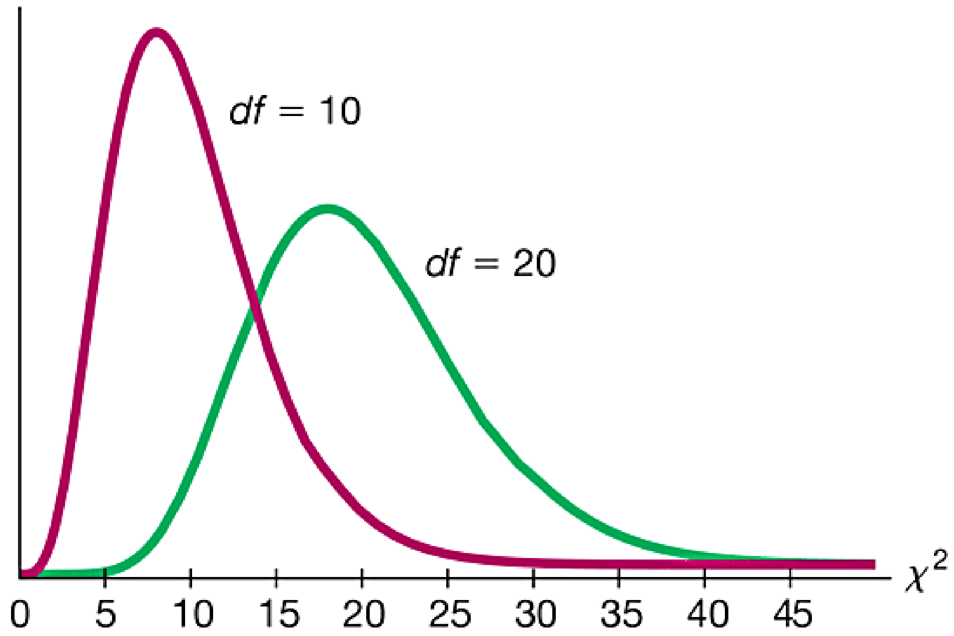
Properties of the Chi-Square Distribution (1 of 3)

1. All values of χ^2 are nonnegative, and the distribution is not symmetric.



Properties of the Chi-Square Distribution (2 of 3)

2. There is a different χ^2 distribution for each number of degrees of freedom.



Properties of the Chi-Square Distribution (3 of 3)

3. The critical values are found in Table A-4 using **degrees of freedom =** $n - 1$

An important note if using Table A-4 for finding critical values:

In Table A-4, each critical value of χ^2 in the body of the table corresponds to an area given in the top row of the table, and each area in that top row is a cumulative area to the right of the critical value.

Example: *P*-Value Method: Do Super Model Heights Vary Less? (1 of 7)

Listed below are the heights (cm) for the simple random sample of female supermodels. Use a 0.01 significance level to test the claim that supermodels have heights with a standard deviation that is less than $\sigma = 7.5$ cm for the population of women. Does it appear that heights of supermodels vary less than heights of women from the population?

178	177	176	174	175	178	175	178	178
	177	180	176	180	178	180	176	

Example: *P*-Value Method: Do Super Model Heights Vary Less? (2 of 7)

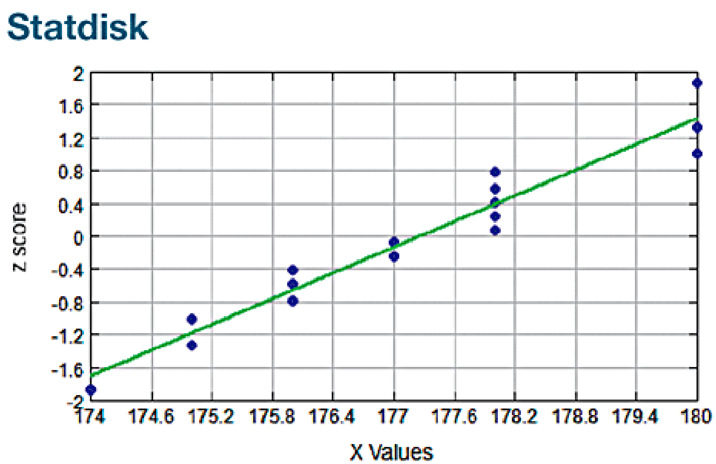
Solution

Requirement Check (1) The sample is a simple random sample.

Example: *P*-Value Method: Do Super Model Heights Vary Less? (3 of 7)

Solution

Requirement Check (2) In checking for normality, we see that the sample has no outliers, the normal quantile plot shows points that are reasonably close to a straight-line pattern, and there is no other pattern that is not a straight line.



Example: *P*-Value Method: Do Super Model Heights Vary Less? (4 of 7)

Solution

Technology capable of conducting this test will typically display the *P*-value. StatCrunch can be used as described at the end of this section, and the result will be as shown in the figure. (Instead of using the assumed value of σ for H_0 and H_1, StatCrunch uses σ^2. For the null hypothesis, $\sigma = 7.5$ is equivalent to $\sigma^2 = 7.52^2 = 56.25$.)

The display shows that the test statistic is $\chi^2 = 0.907$ (rounded) and the *P*-value is less than 0.0001.

StatCrunch

Hypothesis test results:

σ^2 : Variance of population

$H_0 : \sigma^2 = 56.25$

$H_A : \sigma^2 < 56.25$

Variance	Sample Var.	DF	Chi-Square Stat	P-value
σ2	3.4000004	15	0.90666677	<0.0001

Example: *P*-Value Method: Do Super Model Heights Vary Less? (5 of 7)

Solution

Step 1: The claim "the standard deviation is less than 7.5 cm" is expressed as $\sigma < 7.5$ cm.

Step 2: If the original claim is false, then $\sigma \geq 7.5$ cm.

Step 3: The expression $\sigma < 7.5$ cm does not contain equality, so it becomes the alternative hypothesis. The null hypothesis is the statement that $\sigma = 7.5$ cm.

H_0: $\sigma = 7.5$ cm

H_1: $\sigma < 7.5$ cm (original claim)

Example: *P*-Value Method: Do Super Model Heights Vary Less? (6 of 7)

Solution

Step 4: The significance level is $\alpha = 0.01$.

Step 5: Because the claim is made about σ, we use the χ^2 (chi-square) distribution.

Step 6: The StatCrunch display shows the test statistic of $\chi^2 = 0.907$ and it shows that the *P*-value is less than 0.0001.

StatCrunch

Hypothesis test results:

σ^2 : Variance of population

$H_0 : \sigma^2 = 56.25$

$H_A : \sigma^2 < 56.25$

Variance	Sample Var.	DF	Chi-Square Stat	P-value
σ^2	3.4000004	15	0.90666677	<0.0001

Example: *P*-Value Method: Do Super Model Heights Vary Less? (7 of 7)

Solution

Step 7: Because the *P*-value is less than the significance level of $\alpha = 0.01$, we reject H_0.

Interpretation

Step 8: There is sufficient evidence to support the claim that female supermodels have heights with a standard deviation that is less than 7.5 cm for the population of women. It appears that heights of supermodels do vary less than heights of women in the general population.

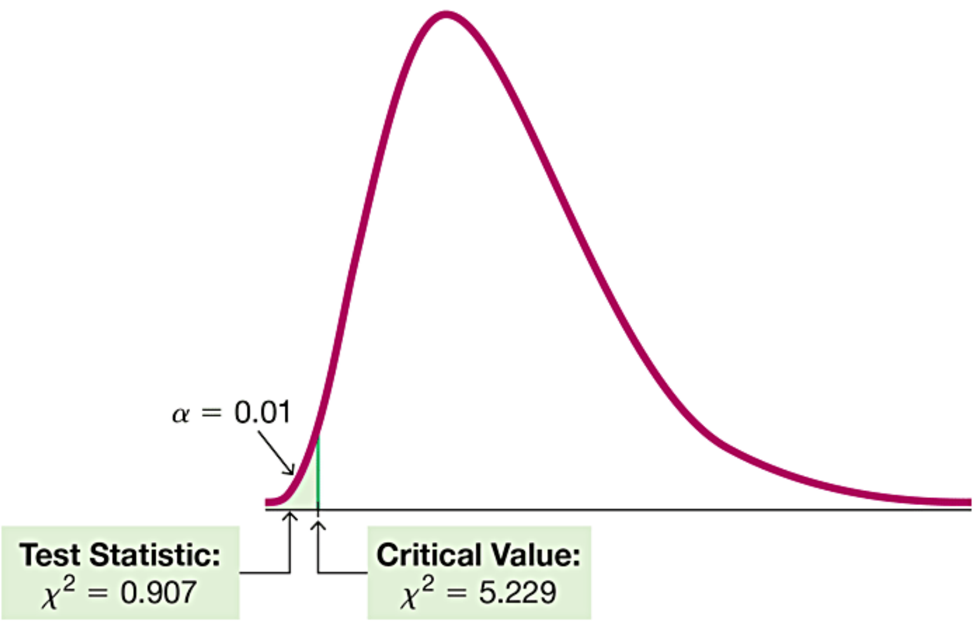
Critical Value Method (1 of 2)

Steps 1 through 5 in the previous example would be the same. In Step 6, the test statistic is calculated by using σ = 7.5 cm (as assumed in the null hypothesis), n = 16, and s = 1.843909 cm, which is the unrounded standard deviation computed from the original list of 16 heights. We get this test statistic:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(16-1)1.843909^2}{7.5^2} = 0.907$$

Critical Value Method (2 of 2)

The critical value of $\chi^2 = 5.229$ is found from Table A-4, and it corresponds to 15 degrees of freedom and an "area to the right" of 0.99 (based on the significance level of 0.01 for a left-tailed test).



Example: Super Model Heights: Critical Value Method (1 of 4)

In Step 7 we reject the null hypothesis because the test statistic of $\chi^2 = 0.907$ falls in the critical region.

We conclude that there is sufficient evidence to support the claim that supermodels have heights with a standard deviation that is less than 7.5 cm for the population of women.

Example: Super Model Heights: Critical Value Method (2 of 4)

Solution

First, we should be careful to select the correct confidence level. Because the hypothesis test is left-tailed and the significance level is 0.01, we should use a confidence level of 98%, or 0.98.

Example: Super Model Heights: Critical Value Method (3 of 4)

Solution

We can use the sample data listed in the example to construct a 98% confidence interval estimate of σ. We use n = 16, s = 1.843909 cm, $\chi^2_L = 5.229$, and $\chi^2_R = 30.578$.

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$1.3 \text{ cm} < \sigma < 3.1 \text{ cm}$$

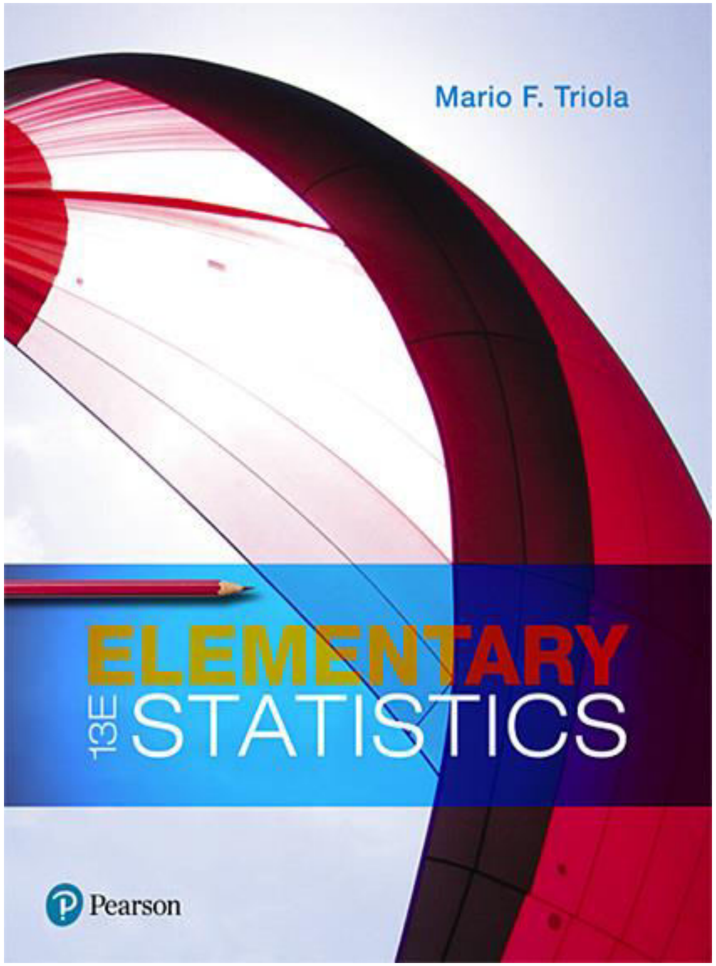
Example: Super Model Heights: Critical Value Method (4 of 4)

Solution

With this confidence interval, we can support the claim that $\sigma < 7.5$ cm because all values of the confidence interval are less than 7.5 cm.

Elementary Statistics

Thirteenth Edition



Chapter 8

Hypothesis Testing

Hypothesis Testing

8-1 Basics of Hypothesis Testing

8-2 Testing a Claim about a Proportion

8-3 Testing a Claim About a Mean

8-4 Testing a Claim About a Standard Deviation or Variance

Key Concept

Testing a claim about a population mean is one of the most important methods presented in this book.

Part 1 of this section deals with the very realistic and commonly used case in which the population standard deviation σ is not known.

Part 2 includes a brief discussion of the procedure used when σ is known, which is very rare.

Testing Claims About a Population Mean with σ Not Known: Objective

Objective

Use a formal hypothesis test to test a claim about a population mean μ.

Testing Claims About a Population Mean with σ Not Known: Notation

Notation

n = sample size

$\bar{x}$ = **sample** mean

s = **sample** standard deviation

$\mu_{\bar{x}}$ = **population** mean

Testing Claims About a Population Mean with σ Not Known: Requirements

Requirements

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

Test Statistic for Testing a Claim About a Mean

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

- ***P*-values:** Use technology or use the Student *t* distribution (Table A-3) with degrees of freedom given by df = $n - 1$.
- **Critical values:** Use the Student *t* distribution (Table A-3) with degrees of freedom given by df = $n - 1$.

Requirement of Normality or $n > 30$

- If the original population is not itself normally distributed, we use the condition $n > 30$ for justifying use of the normal distribution.
- Sample sizes of 15 to 30 are sufficient if the population has a distribution that is not far from normal.
- In this text we use the simplified criterion of $n > 30$ as justification for treating the distribution of sample means as a normal distribution, regardless of how far the distribution departs from a normal distribution.

Important Properties of the Student *t* Distribution

1. The Student *t* distribution is different for different sample sizes.
2. The Student *t* distribution has the same general bell shape as the standard normal distribution; its wider shape reflects the greater variability that is expected when *s* is used to estimate *σ*.
3. The Student *t* distribution has a mean of $t = 0$.
4. The standard deviation of the Student *t* distribution varies with the sample size and is greater than 1.
5. As the sample size *n* gets larger, the Student *t* distribution gets closer to the standard normal distribution.

Example: Adult Sleep (1 of 8)

The author obtained times of sleep for randomly selected adult subjects included in the National Health and Nutrition Examination Study, and those times (hours) are listed below. Here are the unrounded statistics for this sample: $n = 12$, $\bar{x} = 6.83333333$ hours, $s = 1.99240984$ hours.

A common recommendation is that adults should sleep between 7 hours and 9 hours each night. Use the *P*-value method with a 0.05 significance level to test the claim that the mean amount of sleep for adults is less than 7 hours.

4 8 4 4 8 6 9 7 7 10 7 8

Example: Adult Sleep (2 of 8)

Solution

Requirement Check

(1) The sample is a simple random sample.

(2) The second requirement is that "the population is normally distributed or $n > 30$." The sample size is $n = 12$, which does not exceed 30, so we must determine whether the sample data appear to be from a normally distributed population.

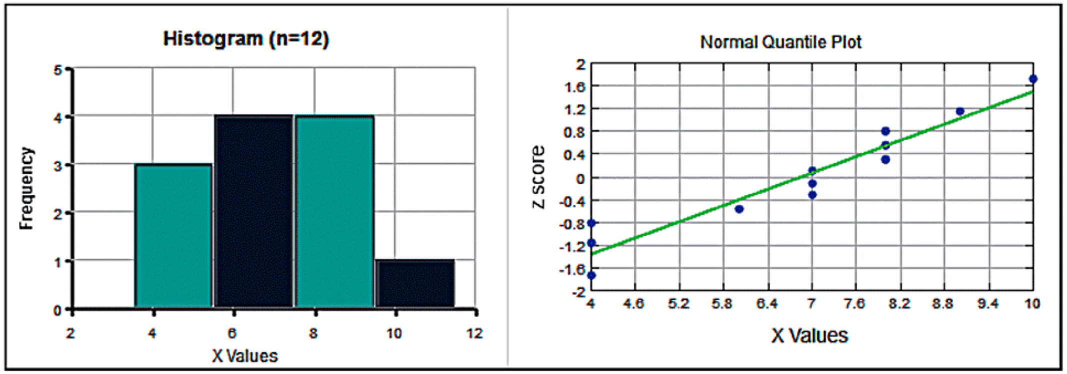
Example: Adult Sleep (3 of 8)

Solution

Requirement Check

The accompanying histogram and normal quantile plot, along with the apparent absence of outliers, indicate that the sample appears to be from a population with a distribution that is approximately normal. Both requirements are satisfied.

Statdisk



Example: Adult Sleep (4 of 8)

Solution

Step 1: The claim that "the mean amount of adult sleep is less than 7 hours" becomes $\mu < 7$ hours.

Step 2: The alternative to the original claim is $\mu \geq 7$ hours.

Step 3: Because the statement $\mu < 7$ hours does not contain the condition of equality, it becomes the alternative hypothesis H_1. The null hypothesis H_0 is the statement that $\mu = 7$ hours.

- H_0: $\mu = 7$ hours (null hypothesis)
- H_1: $\mu < 7$ hours (alternative hypothesis and original claim)

Example: Adult Sleep (5 of 8)

Solution

Step 4: As specified in the statement of the problem, the significance level is $\alpha = 0.05$.

Step 5: Because the claim is made about the **population mean** μ, the sample statistic most relevant to this test is the **sample mean** $\bar{x}$, and we use the t distribution.

Example: Adult Sleep (6 of 8)

Solution

Step 6: The sample statistics of $n = 12$, $\bar{x} = 6.83333333$ hours, $s = 1.99240984$ hours are used to calculate the test statistic as follows, but technologies provide the test statistic of $t = -0.290$. In calculations such as the following, it is good to carry extra decimal places and not round.

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{6.83333333 - 7}{\frac{1.99240984}{\sqrt{12}}} = -0.290$$

Example: Adult Sleep (7 of 8)

Solution

_P_-Value with Technology
We could use technology to obtain the *P*-value. The *P*-value is 0.3887 (rounded).

TI-83/84 Plus

NORMAL FLOAT AUTO REAL RADIAN MP

T-Test
μ<7
t=-.2897748534
p=.3886888459
x̄=6.833333333
Sx=1.99240984
n=12

Statdisk

t Test
Test Statistic, t: -0.2898
Critical t: -1.7959
P-Value: 0.3887

90% Confidence interval:
5.800414 < μ < 7.866252

Excel (XLSTAT)

Difference	-0.1667
t (Observed value)	-0.2898
t (Critical value)	-1.7959
DF	11
p-value (one-tailed)	0.3887
alpha	0.05

Minitab

Test of μ = 7 vs < 7

Variable	N	Mean	StDev	SE Mean	95% Upper Bound	T	P
Sleep	12	6.833	1.992	0.575	7.866	−0.29	0.389

JMP

Hypothesized Value	7
Actual Estimate	6.83333
DF	11
Std Dev	1.99241
t Test	
Test Statistic	-0.2898
Prob > |t|	0.7774
Prob > t	0.6113
Prob < t	0.3887

StatCrunch

Hypothesis test results:
μ : Mean of variable
$H_0 : \mu = 7$
$H_A : \mu < 7$

Variable	Sample Mean	Std. Err.	DF	T-Stat	P-value
Sleep	6.8333333	0.57515918	11	-0.28977485	0.3887

SPSS

	Test Value = 7					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
SLEEP	-.290	11	.777	-.16667	-1.4326	1.0993

Pearson

Example: Adult Sleep (8 of 8)

Solution

Step 7: Because the *P*-value of 0.3887 is greater than the significance level of $\alpha = 0.05$, we fail to reject the null hypothesis.

Interpretation

Step 8: Because we fail to reject the null hypothesis, we conclude that there is not sufficient evidence to support the claim that the mean amount of adult sleep is less than 7 hours.

P-Value Method without Technology

If suitable technology is not available, we can use Table A-3 to identify a **range of values** containing the *P*-value.

In using Table A-3, keep in mind that it is designed for positive values of *t* and right-tail areas only, but left-tail areas correspond to the same *t* values with negative signs.

Example: Adult Sleep: *P*-Value Method without Technology

The previous example is a left-tailed test with a test statistic of $t = -0.290$ and a sample size of $n = 12$, so the number of degrees of freedom is $\text{df} = n - 1 = 11$.

Using the test statistic of $t = -0.290$ with Table A-3, examine the values of t in the row for df = 11 to see that 0.290 is less than all of the listed t values in the row, which indicates that the area in the left tail below the test statistic of $t = -0.290$ is greater than 0.10.

In this case, Table A-3 allows us to conclude that the *P*-value > 0.10, but technology provided the *P*-value of 0.3887. With the *P*-value > 0.10, the conclusions are the same as in the previous example.

Example: Adult Sleep: Critical Value Method (1 of 2)

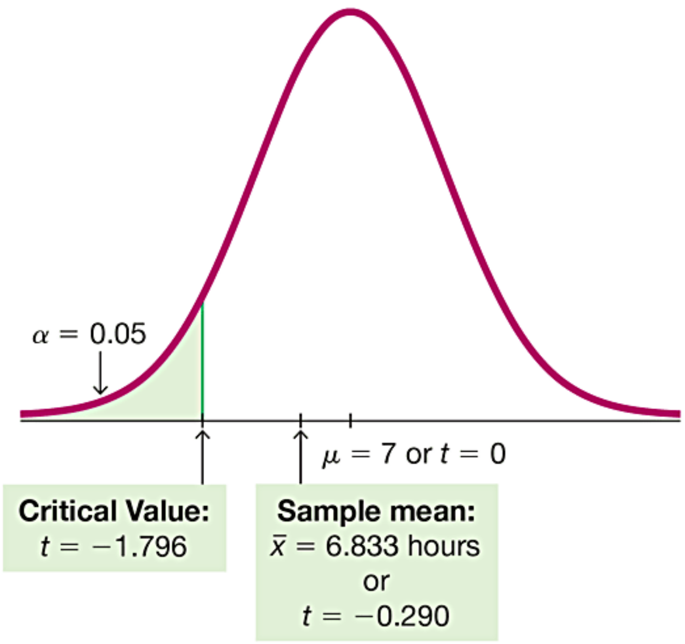
Example 1 is a left-tailed test with test statistic $t = -0.290$. The sample size is $n = 12$, so the number of degrees of freedom is df $= n - 1 = 11$.

Given the significance level of $a = 0.05$, refer to the row of Table A-3 corresponding to 11 degrees of freedom, and refer to the column identifying an “area in one tail” of 0.05. The intersection of the row and column yields the critical value of $t = 1.796$, but this test is left-tailed, so the actual critical value is $t = -1.796$.

Example: Adult Sleep: Critical Value Method (2 of 2)

The figure shows that the test statistic of $t = -0.290$ does not fall within the critical region bounded by the critical value $t = -1.796$, so we fail to reject the null hypothesis.

The conclusions are the same as those given in first example.



Example: Adult Sleep: Confidence Interval Method

The first example is a left-tailed test with significance level $\alpha = 0.05$, so we should use 90% as the confidence level.

For the sample data given in the first example, here is the 90% confidence interval estimate of μ: 5.80 hours $< \mu <$ 7.87 hours.

In testing the claim that $\mu < 7$ hours, we use H_0: $\mu = 7$ hours, but the assumed value of $\mu = 7$ hours is contained within the confidence interval limits, so the confidence interval is telling us that 7 hours could be the value of μ. We don't have sufficient evidence to reject H_0: $\mu = 7$ hours, so we fail to reject this null hypothesis and we get the same conclusions.

Example: Is the Mean Body Temperature Really 98.6°F? (1 of 8)

Data Set 3 "Body Temperatures" in Appendix B includes measured body temperatures with these statistics for 12 AM on day 2: $n = 106$, $\bar{x} = 98.20°\text{F}$, $s = 0.62°\text{F}$. Use a 0.05 significance level to test the common belief that the population mean is 98.6°F.

Example: Is the Mean Body Temperature Really 98.6°F?

Solution

Requirement Check

(1) With the study design used, we can treat the sample as a simple random sample.

(2) The second requirement is that "the population is normally distributed or $n > 30$." The sample size is $n = 106$, so the second requirement is satisfied and there is no need to investigate the normality of the data. Both requirements are satisfied.

Example: Is the Mean Body Temperature Really 98.6°F? (3 of 8)

Solution

Step 1: The claim that "the population mean is 98.6°F" becomes $\mu = 98.6$°F when expressed in symbolic form.

Step 2: The alternative to the original claim is $\mu \neq 98.6$°F.

Step 3: Because the statement $\mu \neq 98.6$°F does not contain the condition of equality, it becomes the alternative hypothesis H_1. The null hypothesis H_0 is the statement that $\mu = 98.6$°F.

H_0: $\mu = 98.6$°F (null hypothesis and original claim)

H_1: $\mu \neq 98.6$°F (alternative hypothesis)

Example: Is the Mean Body Temperature Really 98.6°F? (4 of 8)

Solution

Step 4: As specified in the statement of the problem, the significance level is $\alpha = 0.05$.

Step 5: Because the claim is made about the **population mean** μ, the sample statistic most relevant to this test is the **sample mean** $\bar{x}$. We use the t distribution because the relevant sample statistic is $\bar{x}$ and the requirements for using the t distribution are satisfied.

Example: Is the Mean Body Temperature Really 98.6°F? (5 of 8)

Solution

Step 6: The sample statistics are used to calculate the test statistic as follows, but technologies use unrounded values to provide the test statistic of $t = -6.61$.

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{98.20 - 98.6}{\frac{0.62}{\sqrt{106}}} = -6.64$$

Example: Is the Mean Body Temperature Really 98.6°F?

Solution

***P*-Value** The *P*-value is 0.0000 or 0 + (or "less than 0.01" if using Table A-3).

Critical Values: The critical values are ±1.983 (or ±1.984 if using Table A-3).

Confidence Interval: The 95% confidence interval is $98.08°F < \mu < 98.32°F$.

Example: Is the Mean Body Temperature Really 98.6°F? (7 of 8)

Solution

Step 7: All three approaches lead to the same conclusion: Reject H_0.

- ***P*-Value:** The *P*-value of 0.0000 is less than the significance level of $\alpha = 0.05$.
- **Critical Values:** The test statistic $t = -6.64$ falls in the critical region bounded by ±1.983.
- **Confidence Interval:** The claimed mean of 98.6°F does not fall within the confidence interval of 98.08°F < μ < 98.32°F.

Example: Is the Mean Body Temperature Really 98.6°F? (8 of 8)

Interpretation

Step 8: There is sufficient evidence to warrant **rejection** of the common belief that the population mean is 98.6°F.

Alternative Methods Used When Population Is Not Normal and $n \leq 30$

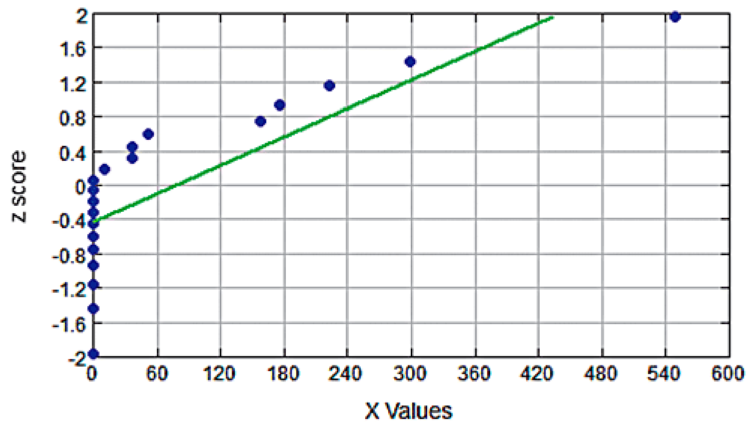
- **Bootstrap Resampling** Use the confidence interval method of testing hypotheses, but obtain the confidence interval using bootstrap resampling. Be careful to use the appropriate confidence level. Reject the null hypothesis if the confidence interval limits do not contain the value of the mean claimed in the null hypothesis.
- **Sign Test** See Section 13-2.
- **Wilcoxon Signed-Ranks Test** See Section 13-3.

Example: Bootstrap Resampling (1 of 3)

Listed below is a random sample of times (seconds) of tobacco use in animated children's movies. Use a 0.05 significance level to test the claim that the sample is from a population with a mean greater than 1 minute, or 60 seconds.

0	223	0	176	0	548	0	37	158		
51	0	0	299	37	0	11	0	0	0	0

Example: Bootstrap Resampling (2 of 3)



Solution

Requirement Check

The *t* test described in Part 1 of this section requires that the population is normally distributed or $n > 30$, but we have $n = 20$ and the accompanying normal quantile plot shows that the sample does not appear to be from a normally distributed population. The *t* test should **not** be used.

Example: Bootstrap Resampling (3 of 3)

Solution

We use the bootstrap resampling method. After obtaining 1000 bootstrap samples and finding the mean of each sample, we sort the means. Because the test is right-tailed with a 0.05 significance level, we use the 1000 sorted sample means to find the 90% confidence interval limits of $P_5 = 29.9$ sec. and $P_{95} = 132.9$ sec. The 90% confidence interval is 29.9 seconds $< \mu <$ 132.9 seconds. Because the assumed mean of 60 seconds is contained within those confidence interval limits, we fail to reject H_0: $\mu = 60$ seconds. There is not sufficient evidence to support H_1: $\mu > 60$ seconds.

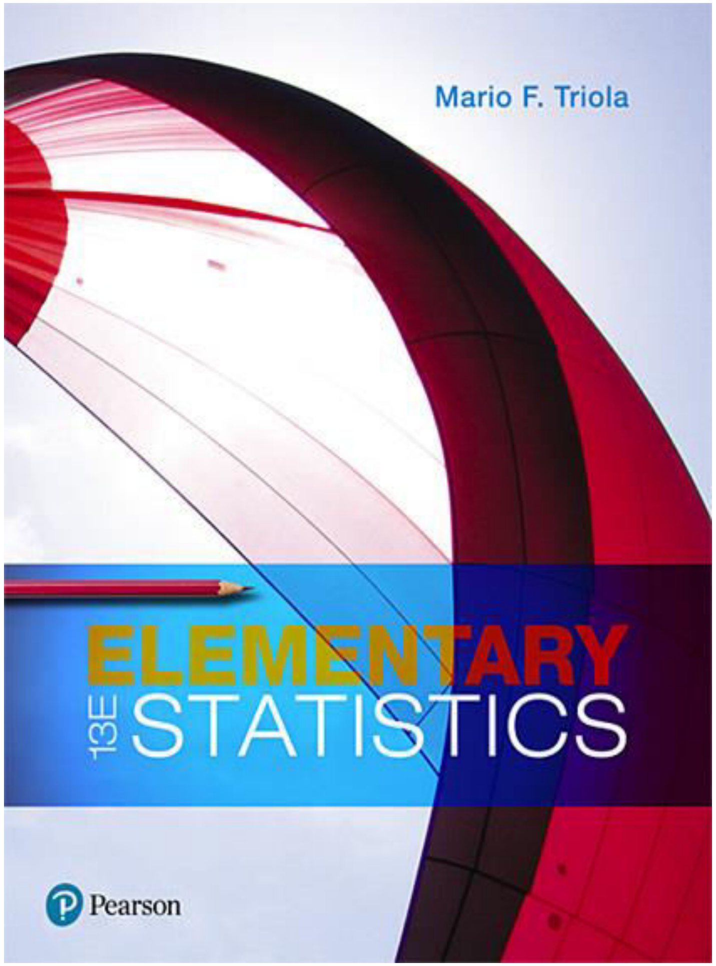
Testing a Claim about a Mean (When σ is Known): Test Statistic

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

- ***P*-value:** Provided by technology, or use the standard normal distribution (Table A-2) with the procedure in Figure 8-3.
- **Critical values:** Use the standard normal distribution (Table A-2).

Elementary Statistics

Thirteenth Edition



Chapter 8

Hypothesis Testing

Hypothesis Testing

8-1 Basics of Hypothesis Testing

8-2 Testing a Claim about a Proportion

8-3 Testing a Claim About a Mean

8-4 Testing a Claim About a Standard Deviation or Variance

Key Concept (1 of 2)

This section describes a complete procedure for testing a claim made about a population proportion p. We illustrate hypothesis testing with the P-value method, the critical value method, and the use of confidence intervals. The methods of this section can be used with claims about population proportions, probabilities, or the decimal equivalents of percentages.

Key Concept (2 of 2)

There are different methods for testing a claim about a population proportion. Part 1 of this section is based on the use of a normal approximation to a binomial distribution, and this method serves well as an introduction to basic concepts, but it is not a method used by professional statisticians. Part 2 discusses other methods that might require the use of technology.

Testing a Claim About a Population Proportion (Normal Approximation Method): Objective

Objective

Conduct a formal hypothesis test of a claim about a population proportion p.

Testing a Claim About a Population Proportion (Normal Approximation Method): Notation

Notation

n = sample size or number of trials

p = population proportion (p is the value used in the statement of the null hypothesis)

$\hat{p} = \frac{x}{n}$ (**sample** proportion)

$q = 1 - p$

Testing a Claim About a Population Proportion (Normal Approximation Method): Requirements (1 of 2)

Requirements

1. The sample observations are a simple random sample.
2. The conditions for a **binomial distribution** are satisfied:
   - There is a fixed number of trials.
   - The trials are independent.
   - Each trial has two categories of “success” and “failure.”
   - The probability of a success remains the same in all trials.

Testing a Claim About a Population Proportion (Normal Approximation Method): Requirements (2 of 2)

Requirements

3. The conditions $np \geq 5$ and $nq \geq 5$ are both satisfied, so **the binomial distribution of sample proportions can be approximated by a normal distribution with** $\mu = np$ and $\sigma = \sqrt{npq}$ (as described in Section 6-6).

   Note that p used here is the **assumed** proportion used in the claim, not the sample proportion $\hat{p}$.

Testing a Claim About a Population Proportion (Normal Approximation Method): Test Statistic for Testing a Claim about a Proportion (1 of 2)

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Testing a Claim About a Population Proportion (Normal Approximation Method): Test Statistic for Testing a Claim about a Proportion (2 of 2)

- ***P*-values:** *P*-values are automatically provided by technology. If technology is not available, use the standard normal distribution (Table A-2) and refer to Figure 8-3 on page 364.
- **Critical values:** Use the standard normal distribution (Table A-2).

Equivalent Methods

When testing claims about proportions, the confidence interval method is not equivalent to the *P*-value and critical value methods, so the confidence interval method could result in a different conclusion. (Both the *P*-value method and the critical value method use the same standard deviation based on the **claimed proportion** *p*, so they are equivalent to each other, but the confidence interval method uses an estimated standard deviation based on the **sample proportion**.)

Recommendation: Use a confidence interval to **estimate** a population proportion, but use the *P*-value method or critical value method for **testing a claim** about a proportion.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (1 of 16)

1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% (or 545) of them responded with "no." Use these results to test the claim that most consumers are uncomfortable with drone deliveries. We interpret "most" to mean "more than half" or "greater than 0.5."

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (2 of 16)

Requirement Check

We first check the three requirements.

1. The 1009 consumers are randomly selected.
2. There is a fixed number (1009) of independent trials with two categories (the subject is uncomfortable with drone deliveries or is not).

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (3 of 16)

Requirement Check

3. The requirements $np \geq 5$ and $nq \geq 5$ are both satisfied with $n = 1009$, $p = 0.5$, and $q = 0.5$. [The value of $p = 0.5$ comes from the claim. We get $np = (1009)(0.5) = 504.5$, which is greater than or equal to 5, and we get $nq = (1009)(0.5) = 504.5$, which is also greater than or equal to 5.]

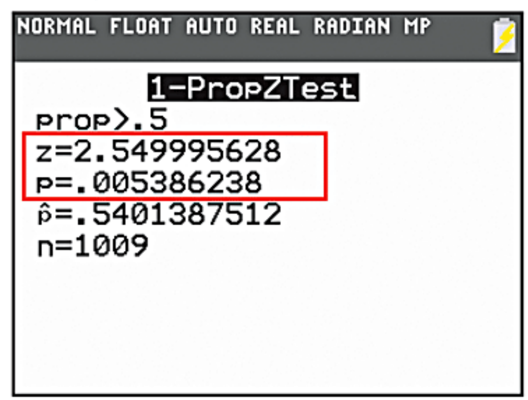
The three requirements are satisfied.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (4 of 16)

Solution: *P*-Value Method

Technology: Computer programs and calculators usually provide a *P*-value, so the *P*-value method is used. See the accompanying TI-83/84 Plus calculator results showing the alternative hypothesis of “prop > 0.5,” the test statistic of $z = 2.55$ (rounded), and the *P*-value of 0.0054 (rounded).

TI-83/84 Plus



Pearson

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (5 of 16)

Solution: *P*-Value Method

Table A-2: If technology is not available, Figure 8-1 on page 360 lists the steps for using the *P*-value method. Using those steps from Figure 8-1, we can test the claim as follows.

Step 1: The original claim is that most consumers are uncomfortable with drone deliveries, and that claim can be expressed in symbolic form as $p > 0.5$.

Step 2: The opposite of the original claim is $p \le 0.5$.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (6 of 16)

Solution: *P*-Value Method

Step 3: Of the preceding two symbolic expressions, the expression $p > 0.5$ does not contain equality, so it becomes the alternative hypothesis. The null hypothesis is the statement that p equals the fixed value of 0.5. We can therefore express H_0 and H_1 as follows:

H_0: $p = 0.5$

H_1: $p > 0.5$ (original claim)

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (7 of 16)

Solution: *P*-Value Method

Step 4: For the significance level, we select $\alpha = 0.05$, which is a very common choice.

Step 5: Because we are testing a claim about a population proportion p, the sample statistic $\hat{p}$ is relevant to this test. The sampling distribution of sample proportions $\hat{p}$ can be approximated by a normal distribution in this case (as described in Section 6-3).

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (8 of 16)

Solution: *P*-Value Method

Step 6: The test statistic $z = 2.55$ can be found by using technology or it can be calculated by using

$\hat{p} = \frac{545}{1009}$ (sample proportion), $n = 1009$ (sample size),

$p = 0.5$ (assumed in the null hypothesis), and

$q = 1 - 0.5 = 0.5.$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{545}{1009} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1009}}} = 2.55$$

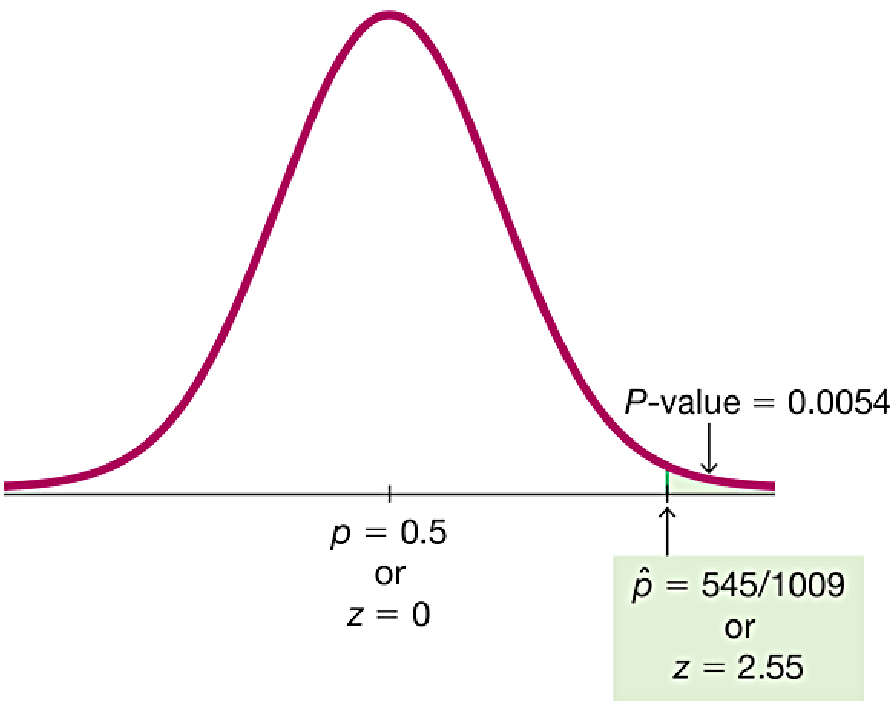
Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (9 of 16)

Solution: *P*-Value Method

Because this hypothesis test is right-tailed with a test statistic of $z = 2.55$. The *P*-value is the area to the right of $z = 2.55$. Referring to Table A-2, the cumulative area to the **left** of $z = 2.55$ is 0.9946, so the area to the right of that test statistic is $1 - 0.9946 = 0.0054$. We get *P*-value = 0.0054.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (10 of 16)

Solution: *P*-Value Method



Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (11 of 16)

Solution: *P*-Value Method

Step 7: Because the *P*-value of 0.0054 is less than or equal to the significance level of $\alpha = 0.05$, we reject the null hypothesis.

Step 8: Because we reject H_0: $p = 0.5$, we support the alternative hypothesis of $p > 0.5$. We conclude that there is sufficient sample evidence to support the claim that more than half of consumers are uncomfortable with drone deliveries.

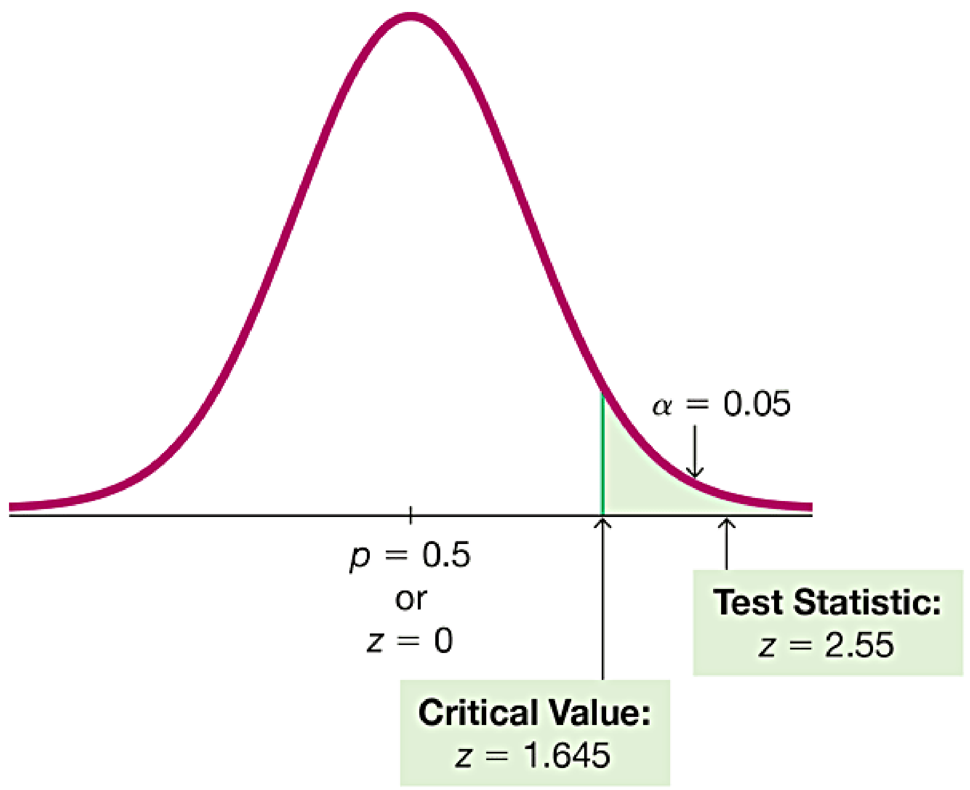
Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (12 of 16)

Solution: Critical Value Method (steps 1 to 5 are the same as the previous method)

Step 6: The test statistic is computed to be $z = 2.55$. With the critical value method, we now find the critical values. This is a right-tailed test, so the area of the critical region is an area of $\alpha = 0.05$ in the right tail. Referring to Table A-2 and applying the methods of Section 6-1, we find that the critical value is $z = 1.645$, which is at the boundary of the critical region.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (13 of 16)

Solution: Critical Value Method



Pearson

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (14 of 16)

Solution: Critical Value Method

Step 7: Because the test statistic does fall within the critical region, we reject the null hypothesis.

Step 8: Because we reject H_0: $p = 0.5$, we conclude that there is sufficient sample evidence to support the claim that most (more than half) consumers are uncomfortable with drone deliveries.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (15 of 16)

Solution: Confidence Interval Method

The claim "The Majority of Consumers Are Not Comfortable with Drone Deliveries," can be tested with a 0.05 significance level by constructing a 90% confidence interval.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (16 of 16)

Solution: Confidence Interval Method

The 90% confidence interval estimate of the population proportion p is found using the sample data of $n = 1009$ and $\hat{p} = \frac{545}{1009}$.

Using the methods of Section 7-1 we get: $0.514 < p < 0.566$. The entire range of values in this confidence interval is greater than 0.5. We are 90% confident that the limits of 0.514 and 0.566 contain the true value of p, the sample data appear to support the claim that most (more than 0.5) consumers are uncomfortable with drone deliveries.

Finding the Number of Successes *x*

When using technology for hypothesis tests of proportions, we must usually enter the sample size n and the number of successes x, but in real applications the sample proportion $\hat{p}$ is often given instead of x. The number of successes x can be found by evaluating $x = n\hat{p}$.

Note that in the next example, the result of 5587.712 adults must be rounded to the nearest whole number of 5588.

Example: Finding the number of Successes *x* (1 of 2)

A study of sleepwalking or “nocturnal wandering” was described in **Neurology** magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked. What is the actual number of adults who have sleepwalked?

Example: Finding the number of Successes *x* (2 of 2)

Solution

The number of adults who have sleepwalked is 29.2% of 19,136, or $0.292 \times 19{,}136 = 5587.712$, but the result must be a whole number, so we round the product to the nearest whole number of 5588.

Example: Fewer Than 30% of Adults have Sleepwalked? (1 of 8)

Using the same sleepwalking data from the previous example (n = 19,136 and $\hat{p}$ = 29.2%), would a reporter be justified in stating that "fewer than 30% of adults have sleepwalked"?

Let's use a 0.05 significance level to test the claim that for the adult population, the proportion of those who have sleepwalked is less than 0.30.

Example: Fewer Than 30% of Adults have Sleepwalked? (2 of 8)

Solution

Requirement Check

(1) The sample is a simple random sample.

(2) There is a fixed number (19,136) of independent trials with two categories (a subject has sleepwalked or has not).

(3) The requirements $np \geq 5$ and $nq \geq 5$ are both satisfied with n = 19,136 and p = 0.30. [We get $np = (19{,}136)(0.30) = 5740.8$, which is greater than or equal to 5, and we also get $nq = (19{,}136)(0.70) = 13{,}395.2$, which is greater than or equal to 5.]

The three requirements are all satisfied.

Example: Fewer Than 30% of Adults have Sleepwalked?

Solution

Step 1: The original claim is expressed in symbolic form as $p < 0.30$.

Step 2: The opposite of the original claim is $p \geq 0.30$.

Step 3: Because $p < 0.30$ does not contain equality, it becomes H_1. We get

H_0: $p = 0.30$ (null hypothesis)

H_1: $p < 0.30$ (alternative hypothesis and original claim)

Example: Fewer Than 30% of Adults have Sleepwalked? (4 of 8)

Solution

Step 4: The significance level is $\alpha = 0.05$.

Step 5: Because the claim involves the proportion p, the statistic relevant to this test is the sample proportion $\hat{p}$ and the sampling distribution of sample proportions can be approximated by the normal distribution.

Example: Fewer Than 30% of Adults have Sleepwalked?

Solution

Step 6: Technology If using technology, the test statistic and the *P*-value will be provided. See the accompanying results from StatCrunch showing that the test statistic is $z = -2.41$ (rounded) and the *P*-value = 0.008.

StatCrunch

Hypothesis test results:
p : Proportion of successes
$H_0 : p = 0.3$
$H_A : p < 0.3$

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	5588	19136	0.29201505	0.0033127149	-2.4103945	0.008

Example: Fewer Than 30% of Adults have Sleepwalked? (6 of 8)

Solution

Table A-2 If technology is not available, proceed as follows to conduct the hypothesis test using the *P*-value method.

The test statistic $z = -2.41$ is calculated as follows:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{5588}{19{,}136} - 0.30}{\sqrt{\frac{(0.30)(0.70)}{19{,}136}}} = -2.41$$

Example: Fewer Than 30% of Adults have Sleepwalked? (7 of 8)

Solution

For this left-tailed test, the *P*-value is the area to the left of the test statistic. Using Table A-2, we see that the area to the left of $z = -2.41$ is 0.0080, so the *P*-value is 0.0080.

Step 7: Because the *P*-value of 0.0080 is less than or equal to the significance level of 0.05, we reject the null hypothesis.

Example: Fewer Than 30% of Adults have Sleepwalked?

Interpretation

Because we reject the null hypothesis, we support the alternative hypothesis. We therefore conclude that there is sufficient evidence to support the claim that fewer than 30% of adults have sleepwalked.