Chapter 9: Testing a Claim

Section 9.2
Tests About a Population Proportion

The Practice of Statistics, 4th edition – For AP*
STARNES, YATES, MOORE

Chapter 9
Testing a Claim

- 9.1 Significance Tests: The Basics
- **9.2 Tests about a Population Proportion**
- 9.3 Tests about a Population Mean

Section 9.2
Tests About a Population Proportion

Learning Objectives

After this section, you should be able to…

- ✓ CHECK conditions for carrying out a test about a population proportion.
- ✓ CONDUCT a significance test about a population proportion.
- ✓ CONSTRUCT a confidence interval to draw a conclusion about for a two-sided test about a population proportion.

Introduction

Confidence intervals and significance tests are based on the sampling distributions of statistics. That is, both use probability to say what would happen if we applied the inference method many times.

Section 9.1 presented the reasoning of significance tests, including the idea of a *P*-value. In this section, we focus on the details of testing a claim about a population proportion.

We'll learn how to perform one-sided and two-sided tests about a population proportion. We'll also see how confidence intervals and two-sided tests are related.

Carrying Out a Significance Test

Recall our basketball player who claimed to be an 80% free-throw shooter. In an SRS of 50 free-throws, he made 32. His sample proportion of made shots, 32/50 = 0.64, is much lower than what he claimed.

Does it provide *convincing evidence* against his claim?

To find out, we must perform a significance test of

$$H_0: p = 0.80$$
$$H_a: p < 0.80$$

where p = the actual proportion of free throws the shooter makes in the long run.

Check Conditions:

In Chapter 8, we introduced three conditions that should be met before we construct a confidence interval for an unknown population proportion: Random, Normal, and Independent. These same three conditions must be verified before carrying out a significance test.

✓ ***Random*** We can view this set of 50 shots as a simple random sample from the population of all possible shots that the player takes.
✓ ***Normal*** Assuming H_0 is true, p = 0.80. then np = (50)(0.80) = 40 and $n(1 - p)$ = (50)(0.20) = 10 are both at least 10, so the normal condition is met.
✓ ***Independent*** In our simulation, the outcome of each shot does is determined by a random number generator, so individual observations are independent.

Alternate Example – Can you be confident of victory?

Problem: Check the conditions for carrying out a significance test to determine if Jack should feel confident of victory in the mayoral election.

Solution: The three required conditions are

Random: A random sample of voters was selected for the poll.

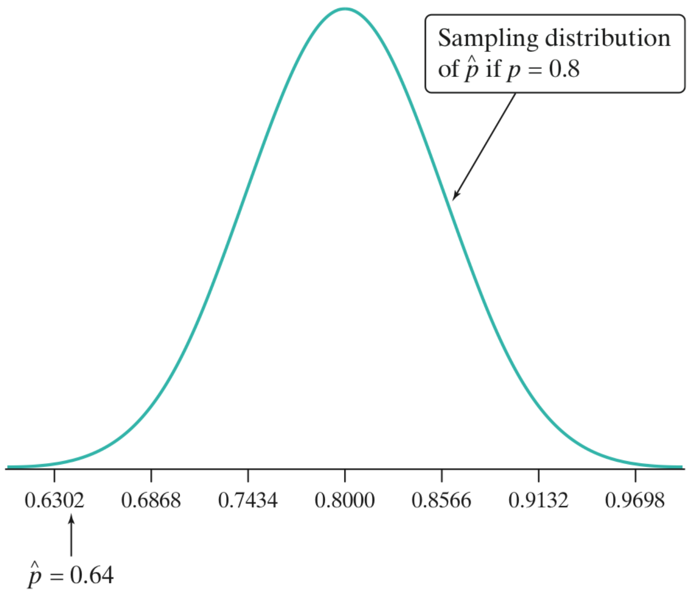
Normal: $np_0 = 100(0.5) = 50 \geq 10$ and $n(1 - p_0) = 100(1 - 0.5) = 50 \geq 10$

Independent: Assuming the poll was done confidentially, one response should not affect other responses. We must assume there are more than 10(100) = 1000 voters since we are sampling without replacement.

Carrying Out a Significance Test

If the null hypothesis $H_0 : p = 0.80$ is true, then the player's sample proportion of made free throws in an SRS of 50 shots would vary according to an approximately Normal sampling distribution with mean

$$\mu_{\hat{p}} = p = 0.80 \text{ and standard deviation } \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.8)(0.2)}{50}} = 0.0566$$



Calculations: Test statistic and *P*-value

A significance test uses sample data to measure the strength of evidence against H_0. Here are some principles that apply to most tests:

- The test compares a statistic calculated from sample data with the value of the parameter stated by the null hypothesis.
- Values of the statistic far from the null parameter value in the direction specified by the alternative hypothesis give evidence against H_0.

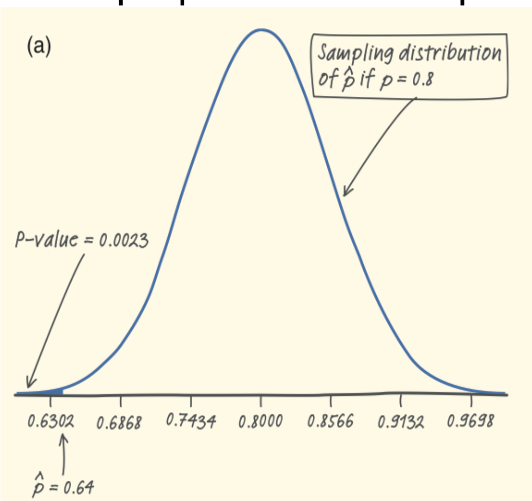
Definition:

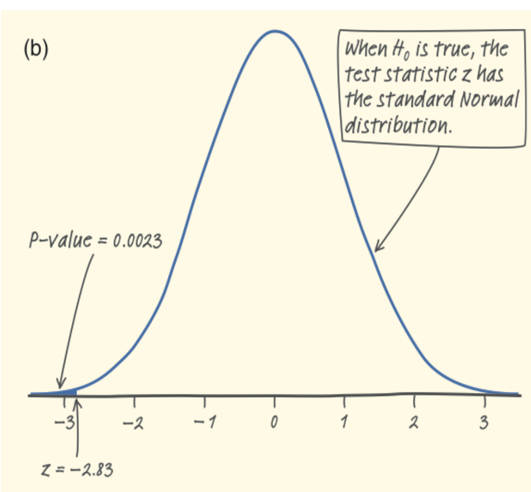
A **test statistic** measures how far a sample statistic diverges from what we would expect if the null hypothesis H_0 were true, in standardized units. That is

$$\text{test statistic} = \frac{\text{statistic - parameter}}{\text{standard deviation of statistic}}$$

Carrying Out a Hypothesis Test

The test statistic says how far the sample result is from the null parameter value, and in what direction, on a standardized scale. You can use the test statistic to find the *P*-value of the test. In our free-throw shooter example, the sample proportion 0.64 is pretty far below the hypothesized value H_0: $p = 0.80$.





$$\text{test statistic} = \frac{\text{statistic - parameter}}{\text{standard deviation of statistic}}$$

$$z = \frac{0.64 - 0.80}{0.0566} = -2.83$$

The shaded area under the curve in (a) shows the *P*-value. (b) shows the corresponding area on the standard Normal curve, which displays the distribution of the *z* test statistic. Using Table A, we find that the *P*-value is $P(z \leq -2.83) = 0.0023$.

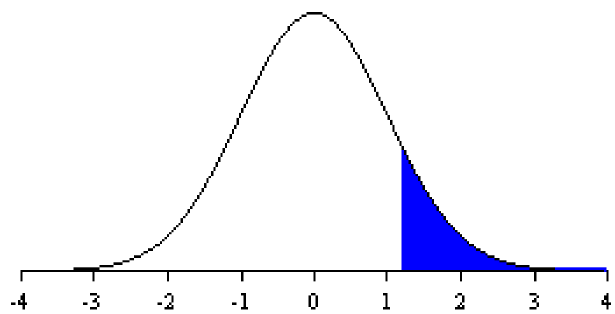
So if H_0 is true, and the player makes 80% of his free throws in the long run, there's only about a 2-in-1000 chance that the player would make as few as 32 of 50 shots.

Alternate Example – How confident can you be?

- **Problem:** In an SRS of 100 voters, 56 favored Jack.
- (a) Calculate the test statistic.
- (b) Find and interpret the *P*-value.
- **Solution:**
- (a) Because the hypothesized value is = 0.5, the standardized test statistic is:

$$z = \frac{0.56 - 0.5}{\sqrt{\frac{05(1-0.5)}{100}}} = 1.20$$

(b) The shaded area under the curve below shows the *P*-value. Using the TI-84, the *P*-value = normalcdf(1.20,100) = 0.1151.



If exactly 50% of all voters support Jack, there is about an 11.5% chance that 56% or more voters would support Jack in a random sample of size 100.

The One-Sample *z* Test for a Proportion

Significance Tests: A Four-Step Process

State: What *hypotheses* do you want to test, and at what significance level? Define any *parameters* you use.

Plan: Choose the appropriate inference *method*. Check *conditions*.

Do: If the conditions are met, perform *calculations*.

- Compute the **test statistic**.
- Find the ***P*-value**.

Conclude: *Interpret* the results of your test in the context of the problem.

When the conditions are met—Random, Normal, and Independent, the sampling distribution of $\hat{p}$ is approximately Normal with mean $\mu_{\hat{p}} = p$ and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.

When performing a significance test, however, the null hypothesis specifies a value for p, which we will call p_0. We assume that this value is correct when performing our calculations.

The One-Sample *z* Test for a Proportion

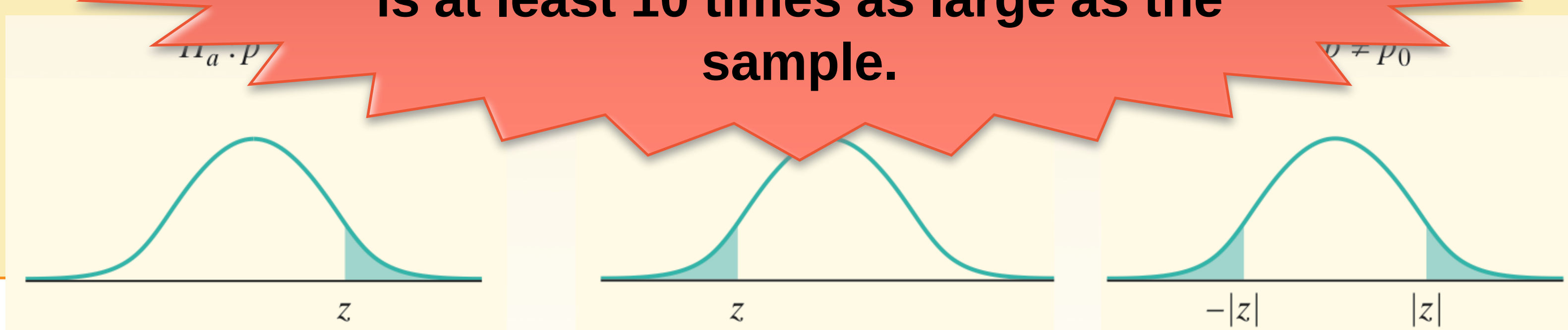
The *z* statistic has approximately the standard Normal distribution when H_0 is true. *P*-values therefore come from the standard Normal distribution. Here is a summary of the details for a **one-sample *z* test for a proportion**.

One-Sample *z* Test for a Proportion

Choose an SRS of size *n* from a large population that contains an unknown proportion *p* of successes. To test the hypothesis $H_0 : p = p_0$, compute the *z* statistic

Find ... large or la...

Use this test only when the expected numbers of successes and failures np_0 and $n(1 - p_0)$ are both at least 10 and the population is at least 10 times as large as the sample.



Example: One Potato, Two Potato

A potato-chip producer has just received a truckload of potatoes from its main supplier. If the producer determines that more than 8% of the potatoes in the shipment have blemishes, the truck will be sent away to get another load from the supplier. A supervisor selects a random sample of 500 potatoes from the truck. An inspection reveals that 47 of the potatoes have blemishes. Carry out a significance test at the α= 0.10 significance level. What should the producer conclude?

State: We want to perform at test at the $\alpha = 0.10$ significance level of

$$H_0: p = 0.08$$
$$H_a: p > 0.08$$

where p is the actual proportion of potatoes in this shipment with blemishes.

Plan: If conditions are met, we should do a one-sample z test for the population proportion p.

✓*Random* The supervisor took a random sample of 500 potatoes from the shipment.

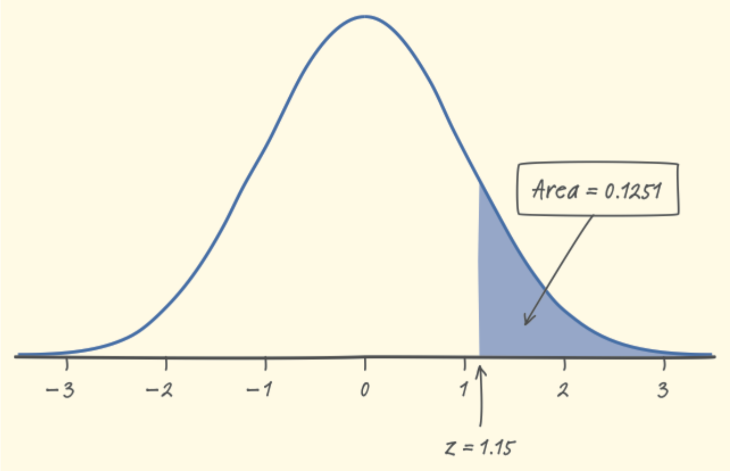
✓*Normal* Assuming $H_0: p = 0.08$ is true, the expected numbers of blemished and unblemished potatoes are $np_0 = 500(0.08) = 40$ and $n(1 - p_0) = 500(0.92) = 460$, respectively. Because both of these values are at least 10, we should be safe doing Normal calculations.

✓*Independent* Because we are sampling without replacement, we need to check the *10% condition*. It seems reasonable to assume that there are at least 10(500) = 5000 potatoes in the shipment.

Example: One Potato, Two Potato

Do: The sample proportion of blemished potatoes is $\hat{p} = 47/500 = 0.094.$

Test statistic $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.094 - 0.08}{\sqrt{\frac{0.08(0.92)}{500}}} = 1.15$$



***P*-value** Using Table A or normalcdf(1.15,100), the desired *P*-value is

$P(z \geq 1.15) = 1 - 0.8749 = 0.1251$

Conclude: Since our *P*-value, 0.1251, is greater than the chosen significance level of $\alpha = 0.10$, we fail to reject H_0. There is not sufficient evidence to conclude that the shipment contains more than 8% blemished potatoes. The producer will use this truckload of potatoes to make potato chips.

Alternate Example: Better to be last?

On shows like American Idol, contestants often wonder if there is an advantage to performing last. To investigate this, a random sample of 600 American Idol fans is selected and they are shown the audition tapes of 12 never-before-seen contestants. For each fan, the order of the 12 videos is randomly determined. Thus, if the order of performance doesn't matter, we would expect approximately 1/12 of the fans to prefer the last contestant they view. In this study, 59 of the 600 fans preferred the last contestant they viewed. Does this data provide convincing evidence that there is an advantage to going last?

State: We want to perform at test at the $\alpha = 0.05$ significance level of

$$H_0: p = 1/12$$
$$H_a: p > 1/12$$

where p = the true proportion of *American Idol* fans who prefer the last performance they see.

Plan: If conditions are met, we will perform a one-sample z test for p.

✓Random: A random sample of American Idol fans was selected and the order in which the videos was viewed was randomized for each subject.

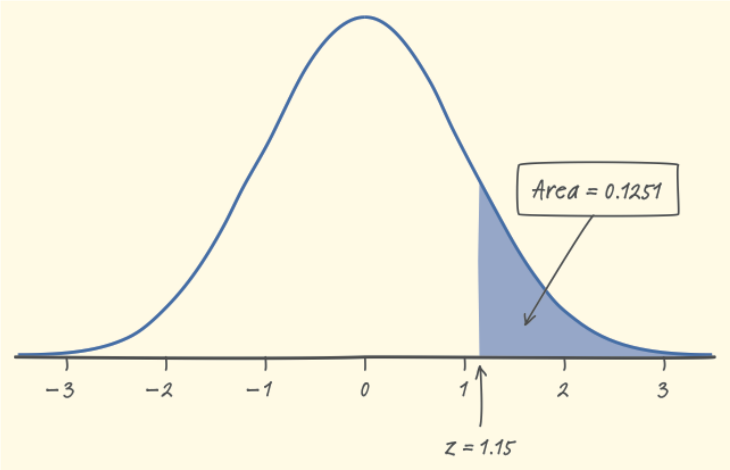
✓Normal: $np_0 = (600)(1/12) = 50 \geq 10$, $n(1 - p_0) = (600)(1 - 1/12) = 550 \geq 10$.

✓Independent: It is reasonable to assume that there are more than 10(600) = 6000 *American Idol* fans.

Alternate Example: Better to be last?

Do: The sample proportion of fans who preferred the last contestant is $\hat{p}$ = 59/600 = 0.098.

Test statistic $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.098 - 0.083}{\sqrt{\frac{0.083(1-0.083)}{600}}} = 1.33$$



***P*-value** Using Table A or normalcdf(1.33,100), the desired *P*-value is $P(z > 1.33) = 0.0918$

Conclude: Since the *P*-value is greater than α (0.0918 > 0.05), we fail to reject the null hypothesis. There is not convincing evidence to conclude that there is an advantage to performing last in *American Idol.*

Two-Sided Tests

According to the Centers for Disease Control and Prevention (CDC) Web site, 50% of high school students have never smoked a cigarette. Taeyeon wonders whether this national result holds true in his large, urban high school. For his AP Statistics class project, Taeyeon surveys an SRS of 150 students from his school. He gets responses from all 150 students, and 90 say that they have never smoked a cigarette. What should Taeyeon conclude? Give appropriate evidence to support your answer.

State: We want to perform at test at the $\alpha = 0.05$ significance level of

$$H_0: p = 0.50$$
$$H_a: p \neq 0.50$$

where p is the actual proportion of students in Taeyeon's school who would say they have never smoked cigarettes.

Plan: If conditions are met, we should do a one-sample z test for the population proportion p.

✓*Random* Taeyeon surveyed an SRS of 150 students from his school.

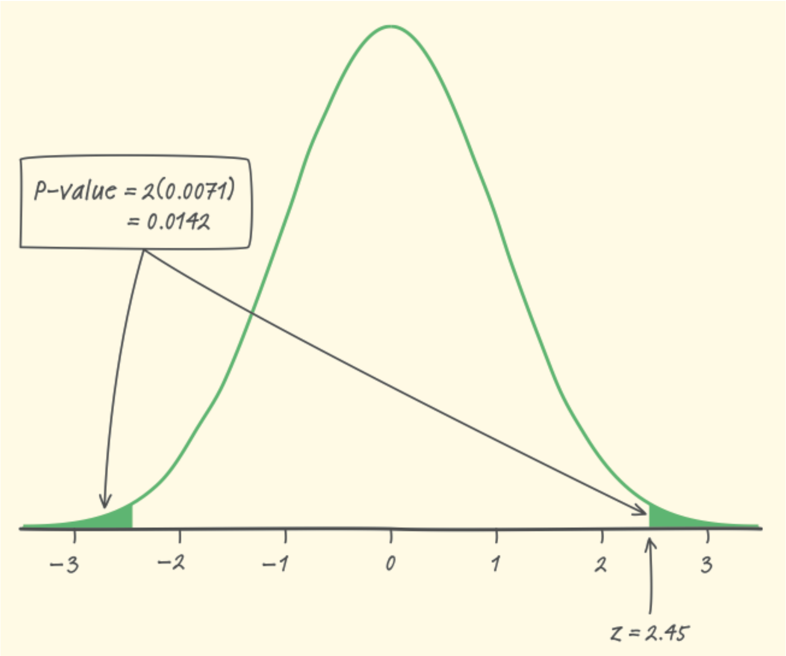
✓*Normal* Assuming H_0: $p = 0.50$ is true, the expected numbers of smokers and nonsmokers in the sample are $np_0 = 150(0.50) = 75$ and $n(1 - p_0) = 150(0.50) = 75$. Because both of these values are at least 10, we should be safe doing Normal calculations.

✓*Independent* We are sampling without replacement, we need to check the *10% condition*. It seems reasonable to assume that there are at least 10(150) = 1500 students a large high school.

Two-Sided Tests

Do: The sample proportion is $\hat{p} = 60/150 = 0.60.$

Test statistic $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.60 - 0.50}{\sqrt{\frac{0.50(0.50)}{150}}} = 2.45$$



***P*-value** To compute this *P*-value, we find the area in one tail and double it. Using Table A or normalcdf(2.45, 100) yields $P(z \geq 2.45) = 0.0071$ (the right-tail area). So the desired *P*-value is 2(0.0071) = 0.0142.

Conclude: Since our *P*-value, 0.0142, is less than the chosen significance level of $\alpha = 0.05$, we have sufficient evidence to reject H_0 and conclude that the proportion of students at Taeyeon's school who say they have never smoked differs from the national result of 0.50.

Alternate Example – Benford's law and fraud

- When the accounting firm AJL and Associates audits a company's financial records for fraud, they often use a test based on Benford's law. Benford's law states that the distribution of first digits in many real-life sources of data is not uniform. In fact, when there is no fraud, about 30.1% of the numbers in financial records begin with the digit 1. However, if the proportion of first digits that are 1 is significantly different from 0.301 in a random sample of records, AJL and Associates does a much more thorough investigation of the company. Suppose that a random sample of 300 expenses from a company's financial records results in only 68 expenses that begin with the digit 1. Should AJL and Associates do a more thorough investigation of this company?

State: We want to perform at test at the $\alpha = 0.05$ significance level of

$$H_0: p = 0.301$$
$$H_a: p \neq 0.301$$

where p = the true proportion of expenses that begin with the digit 1.

Plan: If conditions are met, we will perform a one-sample z test for p.

✓Random: A random sample of expenses was selected.

✓Normal: $np_0 = (300)(0.301) = 90.3 \geq 10$, $n(1 - p_0) = (300)(1 - 0.301) = 209.7 \geq 10$.

✓Independent: It is reasonable to assume that there are more than 10(300) = 3000 expenses in this company's financial records.

Alternate Example – Benford's law and fraud

Do: The sample proportion of expenses that began with the digit 1 is

$$\hat{p} = 68/300 = 0.227.$$

Test statistic $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.227 - 0.301}{\sqrt{\frac{0.301(1-0.301)}{300}}} = -2.79$$

***P*-value** *P*-value: 2*P*(*z* < –2.79) = 2normalcdf(–100,–2.79) = 2(0.0026) = 0.0052

Conclude: Since the *P*-value is less than α (0.0052 < 0.05), we reject the null hypothesis. There is convincing evidence that the proportion of expenses that have first digit of 1 is not 0.301. Therefore, AJL and Associates should do a more thorough investigation of this company.

Why Confidence Intervals Give More Information

The result of a significance test is basically a decision to reject H_0 or fail to reject H_0. When we reject H_0, we're left wondering what the actual proportion *p* might be. A confidence interval might shed some light on this issue.

Taeyeon found that 90 of an SRS of 150 students said that they had never smoked a cigarette. Before we construct a confidence interval for the population proportion *p*, we should check that both the number of successes and failures are at least 10.

✓The number of successes and the number of failures in the sample are 90 and 60, respectively, so we can proceed with calculations.

Our 95% confidence interval is:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.60 \pm 1.96\sqrt{\frac{0.60(0.40)}{150}} = 0.60 \pm 0.078 = (0.522, 0.678)$$

We are 95% confident that the interval from 0.522 to 0.678 captures the true proportion of students at Taeyeon's high school who would say that they have never smoked a cigarette.

Alternate Example – Benford's law and fraud

Problem:

(a) Find and interpret a confidence interval for the true proportion of expenses that begin with the digit 1 for the company in the previous alternate example.

(b) Use your interval from (a) to decide whether this company should be investigated for fraud.

Solution:

(a) *State:* We want to estimate *p* = the true proportion of expenses that begin with the digit 1 at the 95% confidence level.

Plan: We will use a one-sample *z* interval for *p* if the following conditions are satisfied.

Random: A random sample of expenses was selected.

Normal: $n\hat{p} = 68 \geq 10$ and $n(1-\hat{p}) = 232 \geq 10$

Independent: It is reasonable to assume that there are more than 10(300) = 3000 expenses in this company's financial records.

Plan:

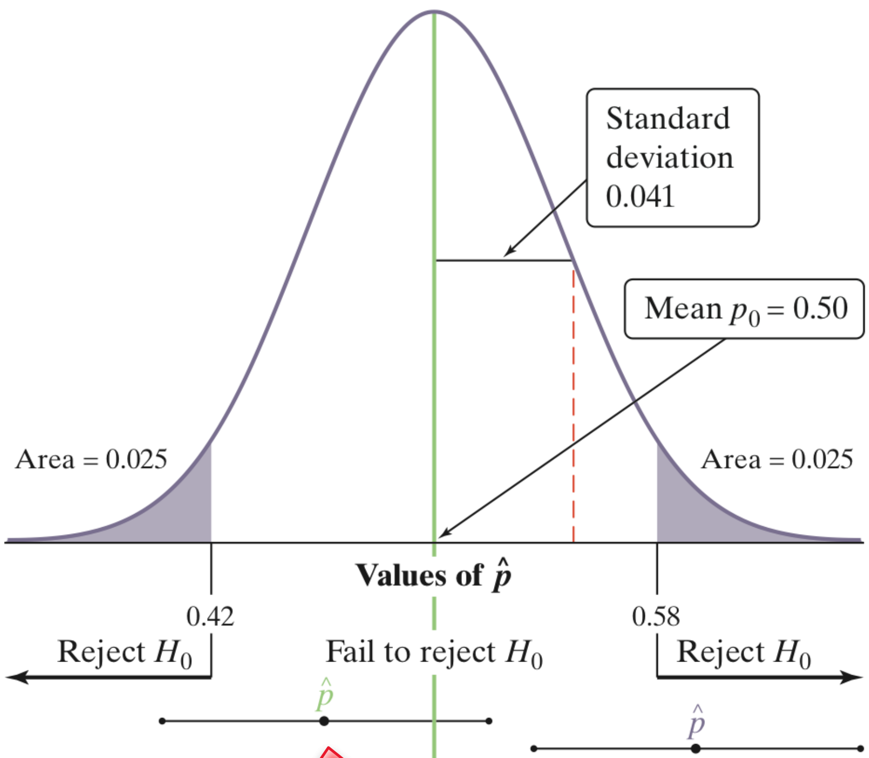
$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.227 \pm 1.96\sqrt{\frac{0.227(1-0.227)}{300}} = 0.227 \pm 0.047 = (0.180, 0.274)$$

Conclude: We are 95% confident that the interval from 0.176 to 0.270 captures the true proportion of expenses at this company that begin with the digit 1.

(b) Since 0.301 is not in the interval from (a), 0.301 is not a plausible value for the true proportion of expenses that begin with the digit 1. Thus, this company should be investigated for fraud.

Confidence Intervals and Two-Sided Tests

There is a link between confidence intervals and two-sided tests. The 95% confidence interval gives an approximate range of p_0's that would not be rejected by a two-sided test at the $\alpha = 0.05$ significance level. The link isn't perfect because the standard error used for the confidence interval is based on the sample proportion, while the denominator of the test statistic is based on the value p_0 from the null hypothesis.



✓ A two-sided test at significance level α (say, $\alpha = 0.05$) and a $100(1 - \alpha)\%$ confidence interval (a 95% confidence interval if $\alpha = 0.05$) give similar information about the population parameter.

✓ However, if the sample proportion falls in the "reject H_0" region, the resulting 95% confidence interval would not include p_0. In that case, both the significance test and the confidence interval would provide evidence that p_0 is not the parameter value.

parameter value.

Section 9.2
Tests About a Population Proportion

Summary

In this section, we learned that…

- ✓ As with confidence intervals, you should verify that the three conditions—Random, Normal, and Independent—are met before you carry out a significance test.

- ✓ Significance tests for $H_0 : p = p_0$ are based on the test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

  with P-values calculated from the standard Normal distribution.

- ✓ The one-sample z test for a proportion is approximately correct when

  (1) the data were produced by random sampling or random assignment;

  (2) the population is at least 10 times as large as the sample; and

  (3) the sample is large enough to satisfy $np_0 \geq 10$ and $n(1 - p_0) \geq 10$ (that is, the expected numbers of successes and failures are both at least 10).

+

Section 9.2
Tests About a Population Proportion

Summary

In this section, we learned that…

- ✓ Follow the four-step process when you carry out a significance test:

  STATE: What hypotheses do you want to test, and at what significance level? Define any parameters you use.

  PLAN: Choose the appropriate inference method. Check conditions.

  DO: If the conditions are met, perform calculations.

  - Compute the test statistic.
  - Find the P-value.

  CONCLUDE: Interpret the results of your test in the context of the problem.

- ✓ Confidence intervals provide additional information that significance tests do not—namely, a range of plausible values for the true population parameter p. A two-sided test of $H_0 : p = p_0$ at significance level α gives roughly the same conclusion as a $100(1 - \alpha)\%$ confidence interval.

Looking Ahead…

In the next Section…

We'll learn how to test a claim about a population mean.

We'll learn about

- ✓ **Carrying out a significance test**
- ✓ **The one-sample *t* test for a mean**
- ✓ **Two-sided tests and confidence intervals**
- ✓ **Paired data and one-sample *t* procedures**