Μηδενική (null) και **Εναλλακτική** (alternative) υπόθεση

DEFINITION 9.1 Null and Alternative Hypotheses

Null hypothesis: A hypothesis to be tested. We use the symbol H_0 to represent the null hypothesis.

Alternative hypothesis: A hypothesis to be considered as an alternative to the null hypothesis. We use the symbol H_a to represent the alternative hypothesis.

null hypothesis always takes the form $\mu = \mu_0$, where μ_0 is some number. In other words, an equals sign (=) should appear in the null hypothesis. We can therefore express the null hypothesis concisely as

$$H_0\colon \mu = \mu_0.$$

- If the primary concern is deciding whether a population mean, μ, is *different from* a specified value μ_0, the alternative hypothesis should be $\mu \neq \mu_0$. In other words, a does-not-equal sign ($\neq$) should appear in the alternative hypothesis. We express such an alternative hypothesis as

$$H_a\colon \mu \neq \mu_0.$$

A hypothesis test whose alternative hypothesis has this form is called a **two-tailed test.**

- If the primary concern is deciding whether a population mean, μ, is *less than* a specified value μ_0, the alternative hypothesis should be $\mu < \mu_0$. In other words, a less-than sign ($<$) should appear in the alternative hypothesis. We express such an alternative hypothesis as

  $$H_a\colon \mu < \mu_0.$$

  A hypothesis test whose alternative hypothesis has this form is called a **left-tailed test.**

- If the primary concern is deciding whether a population mean, μ, is *greater than* a specified value μ_0, the alternative hypothesis should be $\mu > \mu_0$. In other words, a greater-than sign ($>$) should appear in the alternative hypothesis. We express such an alternative hypothesis as

$$H_a\colon \mu > \mu_0.$$

A hypothesis test whose alternative hypothesis has this form is called a **right-tailed test.**

Quality Assurance A company that produces snack foods uses a machine to package 454 g bags of pretzels. We assume that the net weights are normally distributed and that the population standard deviation of all such weights is 7.8 g.[1] A random sample of 25 bags of pretzels has the net weights, in grams, displayed in Table 9.1.

Η **τυπική απόκλιση** του **πληθυσμού** είναι γνωστή!

TABLE 9.1

Weights, in grams, of 25 randomly selected bags of pretzels

465	456	438	454	447
449	442	449	446	447
468	433	454	463	450
446	447	456	452	444
447	456	456	435	450

Statistix 9.0 pretzels_12May09, 12-May-09, 12:34:02 PM

Descriptive Statistics

	WEIGHT
N	25
Lo 95% CI	446.50
Mean	450.00 μέση τιμή
Up 95% CI	453.50
SD	8.4804 τυπική απόκλιση
SE Mean	1.6961
Minimum	433.00
Maximum	468.00

The null and alternative hypotheses for the hypothesis test, as stated in Example 9.1, are

H_0: $\mu = 454$ g (the packaging machine is working properly)

H_a: $\mu \neq 454$ g (the packaging machine is not working properly).

Δειγματοληπτική κατανομή του μέσου όρου

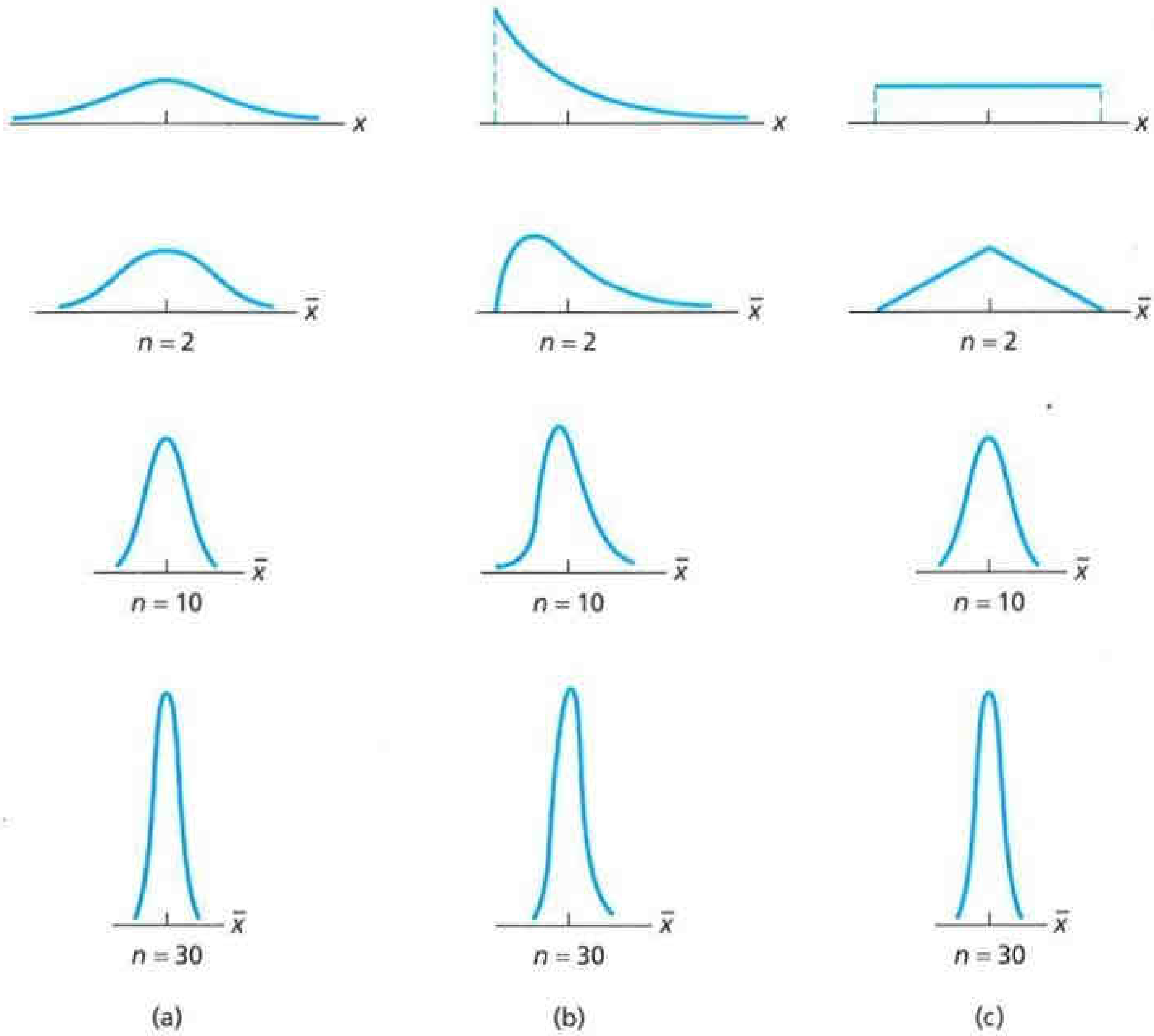
Key Fact 7.4 Sampling Distribution of the Sample Mean

Suppose that a variable x of a population has mean μ and standard deviation σ. Then, for samples of size n,

- the mean of $\bar{x}$ equals the population mean, or $\mu_{\bar{x}} = \mu$;
- the standard deviation of $\bar{x}$ equals the population standard deviation divided by the square root of the sample size, or $\sigma_{\bar{x}} = \sigma/\sqrt{n}$;
- if x is normally distributed, so is $\bar{x}$, regardless of sample size; and
- if the sample size is large, $\bar{x}$ is approximately normally distributed, regardless of the distribution of x.

Thus, if either the variable under consideration is normally distributed or the sample size is large, the distribution of all possible sample means is, at least approximately, a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$.

FIGURE 7.6
Sampling distributions for (a) normal, (b) reverse-J-shaped, and (c) uniform variables



Because $n = 25$, $\sigma = 7.8$, and the weights are normally distributed, Key Fact 7.4 on page 316 implies that

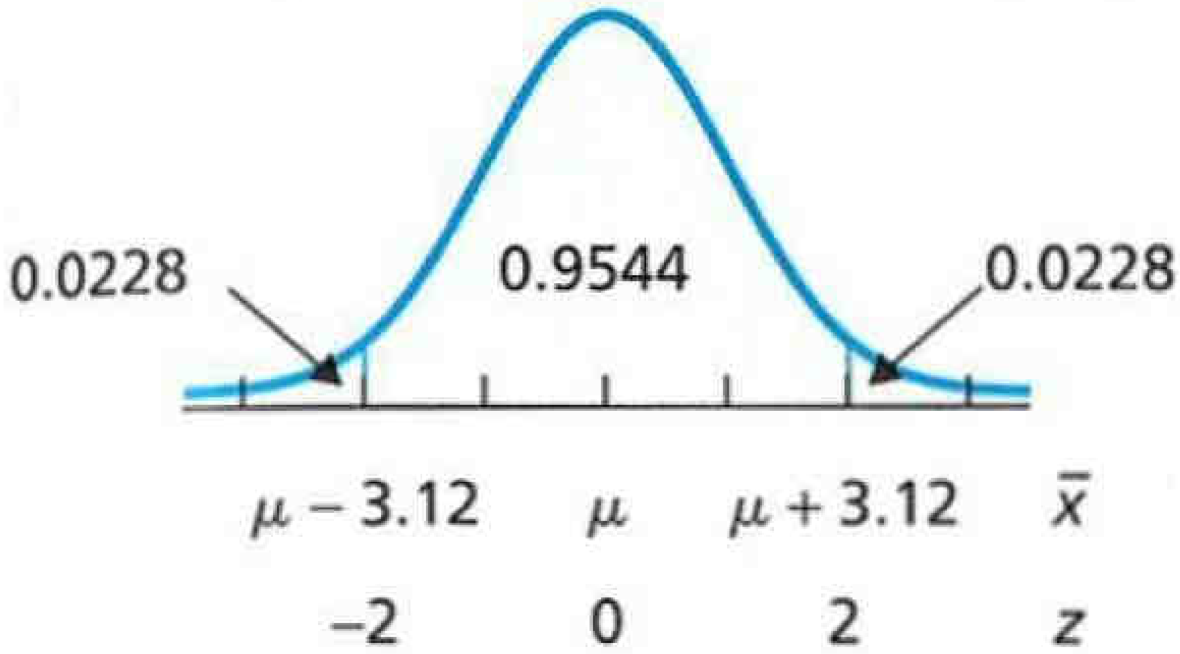
- $\mu_{\bar{x}} = \mu$ (which we don't know),
- $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 7.8/\sqrt{25} = 1.56$, and
- $\bar{x}$ is normally distributed.

In other words, for samples of size 25, the variable $\bar{x}$ is normally distributed with mean μ and a standard deviation of 1.56 g.

The "95.44" part of the 68.26-95.44-99.74 rule states that, for a normally distributed variable, 95.44% of all possible observations lie within two standard deviations to either side of the mean. Applying this part of the rule to the variable $\bar{x}$ and referring to part (c), we see that 95.44% of all samples of 25 bags of pretzels have mean weights within $2 \cdot 1.56 = 3.12$ g of μ. Or, equivalently, only 4.56% of all samples of 25 bags of pretzels have mean weights that are not within 3.12 g of μ, as illustrated in Fig. 9.1.

FIGURE 9.1

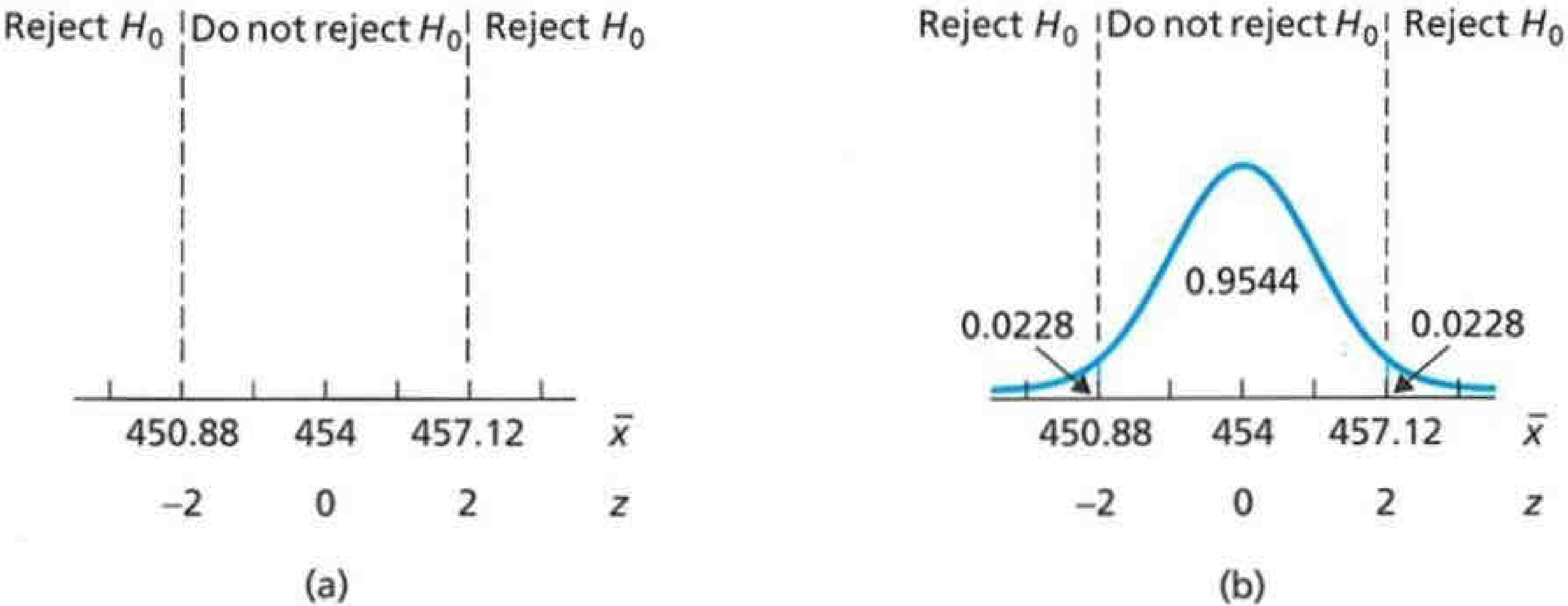
95.44% of all samples of 25 bags of pretzels have mean weights within two standard deviations (3.12 g) of μ



In summary, then, we have obtained the following precise criterion for deciding whether to reject the null hypothesis:

> If the mean weight, $\bar{x}$, of the 25 bags of pretzels sampled is more than two standard deviations from 454 g, reject the null hypothesis, $\mu = 454$ g, and conclude that the alternative hypothesis, $\mu \neq 454$ g, is true. Otherwise, do not reject the null hypothesis.

This criterion is portrayed graphically in Fig. 9.2(a). If the null hypothesis is true, the normal curve associated with $\bar{x}$ is the one with parameters 454 and 1.56; that normal curve is superimposed on Fig. 9.2(a) in Fig. 9.2(b).



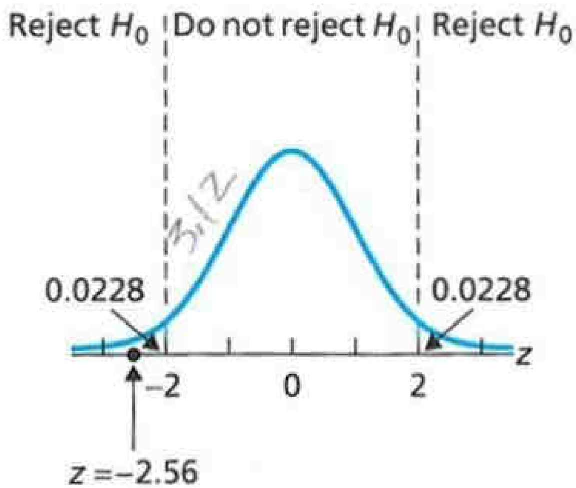
Τυπικό σκορ (standard score)

The mean weight, $\bar{x}$, of the sample of 25 bags of pretzels whose weights are given in Table 9.1 is 450 g. Therefore

$$z = \frac{\bar{x} - 454}{1.56} = \frac{450 - 454}{1.56} = -2.56.$$

That is, the sample mean of 450 g is 2.56 standard deviations below the null-hypothesis population mean of 454 g, as shown in Fig. 9.3.

FIGURE 9.3 Graph showing the number of standard deviations that the sample mean of 450 g is from the null-hypothesis population mean of 454 g



What Does it Mean?

The data provide sufficient evidence to conclude that the packaging machine is not working properly.

Because the mean weight of the 25 bags of pretzels sampled is more than two standard deviations from 454 g, we reject the null hypothesis, $\mu = 454$ g, and conclude that the alternative hypothesis, $\mu \neq 454$ g, is true. ◆

To introduce some additional terminology used in hypothesis testing, we refer to the pretzel packaging hypothesis test of Example 9.4 on page 374. Recall that the null and alternative hypotheses for that hypothesis test are

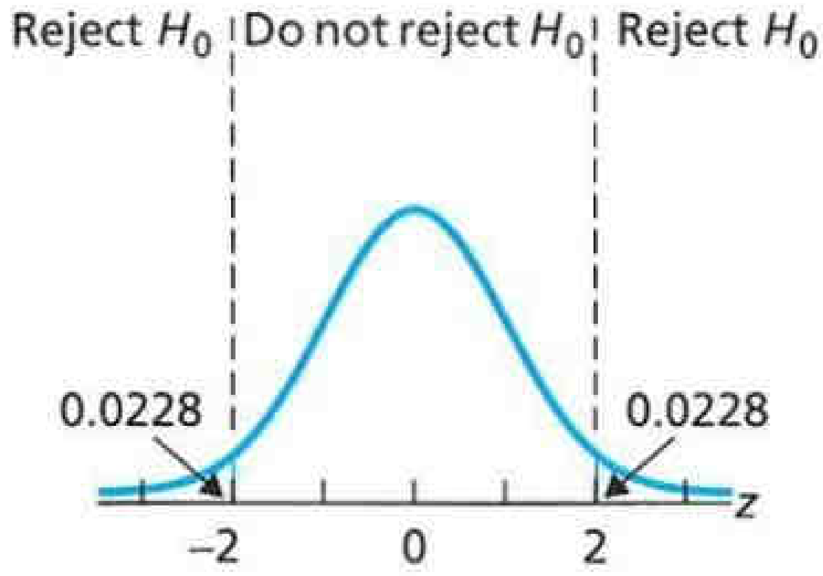
$$H_0: \mu = 454 \text{ g (the packaging machine is working properly)}$$
$$H_a: \mu \neq 454 \text{ g (the packaging machine is not working properly),}$$

where μ is the mean net weight of all bags of pretzels packaged.

As a basis for deciding whether to reject the null hypothesis, in part (e) of Example 9.4 we utilized the variable

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 454}{1.56},$$

which tells us how many standard deviations the sample mean is from the null hypothesis population mean of 454 g. That variable is called the **test statistic** for the hypothesis test.



The set of values for the test statistic that leads us to reject the null hypothesis is called the **rejection region.** In this case, the rejection region consists of all z-scores that lie either to the left of -2 or to the right of 2—that part of the horizontal axis under the shaded areas in Fig. 9.4.

The set of values for the test statistic that leads us not to reject the null hypothesis is called the **nonrejection region,** or **acceptance region.** In this case, the nonrejection region consists of all z-scores that lie between -2 and 2—that part of the horizontal axis under the unshaded area in Fig. 9.4.

The values of the test statistic that separate the rejection and nonrejection regions are called the **critical values.** In this case, the critical values are $z = \pm 2$, as shown in Fig. 9.4. We summarize the preceding discussion in Fig. 9.5.

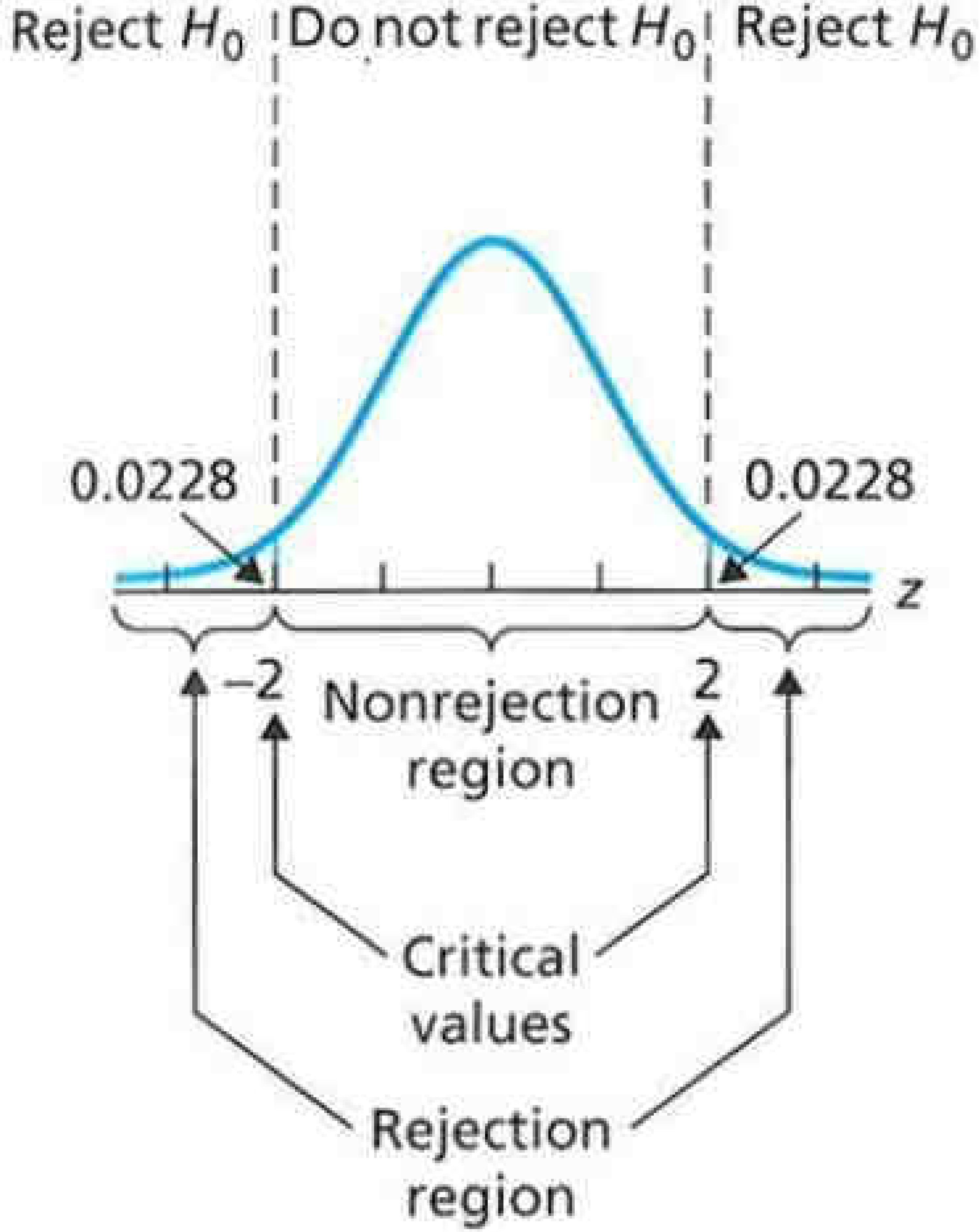
Test Statistic, Rejection Region, Nonrejection Region, Critical Values

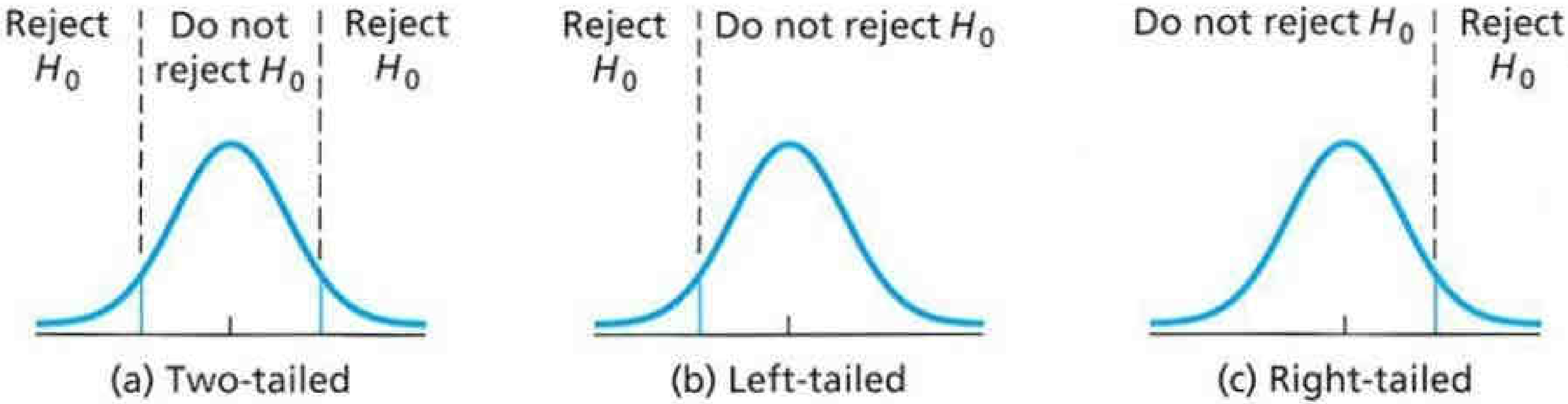
Test statistic: The statistic used as a basis for deciding whether the null hypothesis should be rejected.

Rejection region: The set of values for the test statistic that leads to rejection of the null hypothesis.

Nonrejection region: The set of values for the test statistic that leads to nonrejection of the null hypothesis.

Critical values: The values of the test statistic that separate the rejection and nonrejection regions. A critical value is considered part of the rejection region.


Reject H_0
Do not reject H_0
Reject H_0
0.0228
0.0228
z
−2
2
Nonrejection region
Critical values
Rejection region



	Two-tailed test	Left-tailed test	Right-tailed test
Sign in H_a	$\neq$	$<$	$>$
Rejection region	Both sides	Left side	Right side

There are two types of incorrect decisions—**Type I error,** or rejection of a true null hypothesis, and **Type II error,** or nonrejection of a false null hypothesis, as indicated in Table 9.3 and the following definition.

		H_0 **is:**	
		True	False
Decision:	Do not reject H_0	Correct decision	Type II error
	Reject H_0	Type I error	Correct decision

Type I and Type II Errors

Type I error: Rejecting the null hypothesis when it is in fact true.

Type II error: Not rejecting the null hypothesis when it is in fact false.

Significance Level

The probability of making a Type I error, that is, of rejecting a true null hypothesis, is called the ***significance level, α,*** of a hypothesis test.

A Type II error occurs if the test statistic falls in the nonrejection region when in fact the null hypothesis is false. The probability of that happening, the **Type II error probability,** is denoted β (the Greek letter beta). It depends on the true value of μ.

A concept closely related to that of Type II error probability is *power.* The **power** of a hypothesis test is the probability of not making a Type II error, that is, the probability of rejecting a false null hypothesis. We have

$$\text{Power} = 1 - P(\text{Type II error}) = 1 - \beta.$$

Relation Between Type I and Type II Error Probabilities

For a fixed sample size, the smaller we specify the significance level, α, the larger will be the probability, β, of not rejecting a false null hypothesis.

Possible Conclusions for a Hypothesis Test

Suppose that a hypothesis test is conducted at a small significance level.

- If the null hypothesis is rejected, we conclude that the alternative hypothesis is true.
- If the null hypothesis is not rejected, we conclude that the data do not provide sufficient evidence to support the alternative hypothesis.

Προσοχή: ΠΟΤΕ δεν λέμε "accept"!

Obtaining Critical Values

Suppose that a hypothesis test is to be performed at a specified significance level, α. Then the critical value(s) must be chosen so that, if the null hypothesis is true, the probability is α that the test statistic will fall in the rejection region.

FIGURE 9.7
Critical value(s) for a hypothesis test at the 5% significance level if the test is (a) two-tailed, (b) left-tailed, or (c) right-tailed

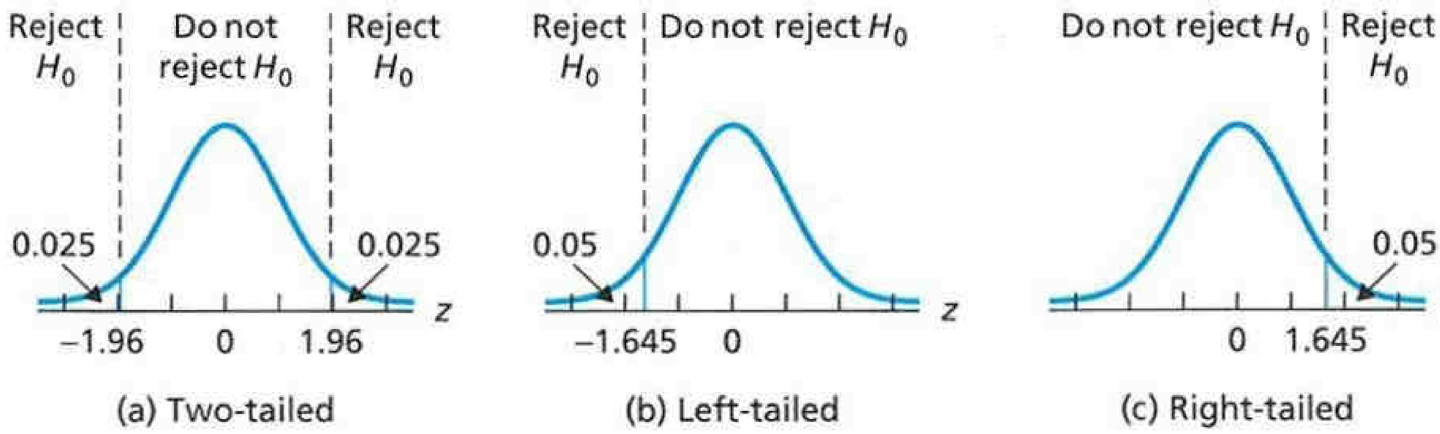
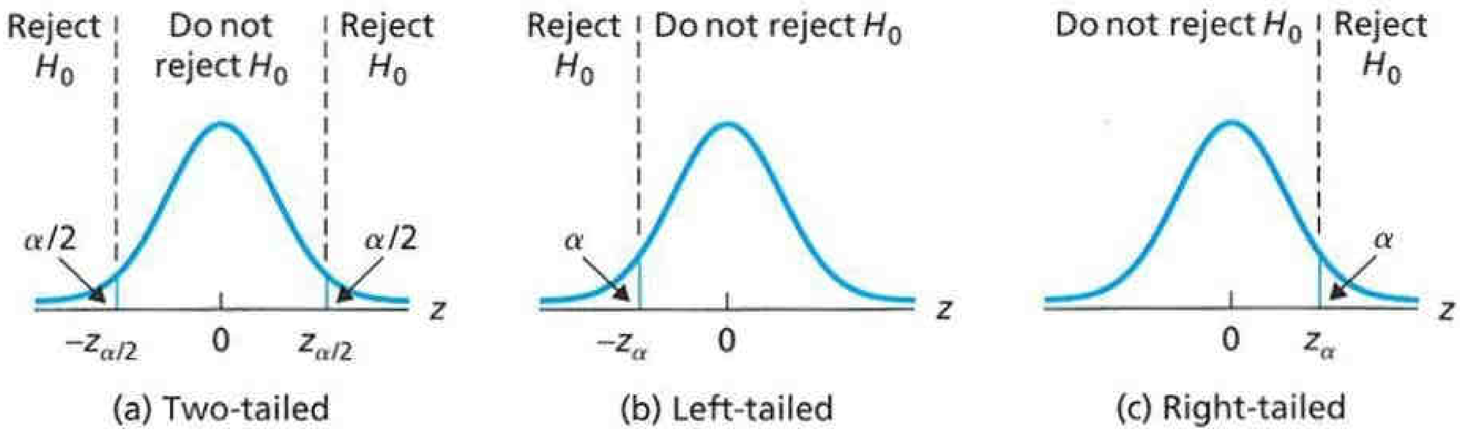


FIGURE 9.8
Critical value(s) for a hypothesis test at the significance level α if the test is (a) two-tailed, (b) left-tailed, or (c) right-tailed



The most commonly used significance levels are 0.10, 0.05, and 0.01. If we consider both one-tailed and two-tailed tests, these three significance levels give rise to five "tail areas." Using the standard-normal table, Table II, we obtained the value of z_α corresponding to each of those five tail areas, as displayed in Table 9.4.

TABLE 9.4
Some important values of z_α

$z_{0.10}$	$z_{0.05}$	$z_{0.025}$	$z_{0.01}$	$z_{0.005}$
1.28	1.645	1.96	2.33	2.575

Procedure 9.1

The One-Sample z-Test for a Population Mean (Critical-Value Approach)

Assumptions

1. Normal population or large sample
2. σ known

Step 1 The null hypothesis is H_0: $\mu = \mu_0$, and the alternative hypothesis is

H_a: $\mu \neq \mu_0$		H_a: $\mu < \mu_0$		H_a: $\mu > \mu_0$
(Two-tailed)	or	(Left-tailed)	or	(Right-tailed)

Step 2 Decide on the significance level, α.

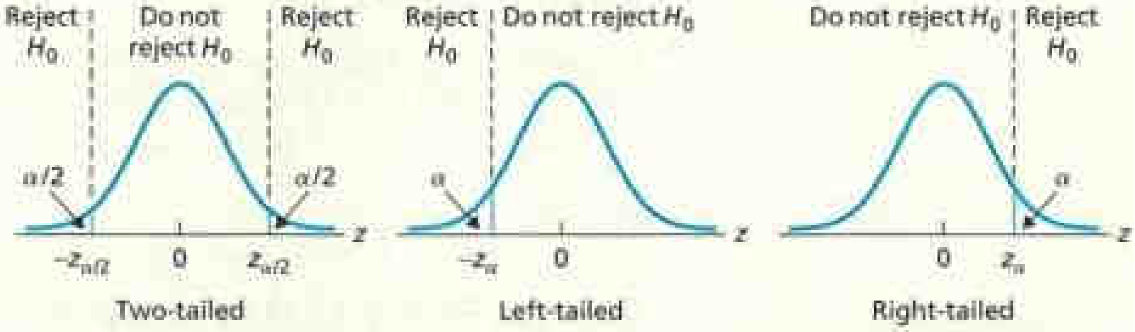
Step 3 Compute the value of the test statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.$$

Step 4 The critical value(s) are

$\pm z_{\alpha/2}$		$-z_\alpha$		z_α
(Two-tailed)	or	(Left-tailed)	or	(Right-tailed)

Use Table II to find the critical value(s).



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.

The One-Sample z-Test for a Population Mean (Critical-Value Approach)

Assumptions

1. Normal population or large sample
2. σ known

Step 1 **The null hypothesis is H_0: $\mu = \mu_0$, and the alternative hypothesis is**

H_a: $\mu \neq \mu_0$	or	H_a: $\mu < \mu_0$	or	H_a: $\mu > \mu_0$
(Two-tailed)		(Left-tailed)		(Right-tailed)

Step 2 Decide on the significance level, α.

Step 3 Compute the value of the test statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.$$

Step 4 The critical value(s) are

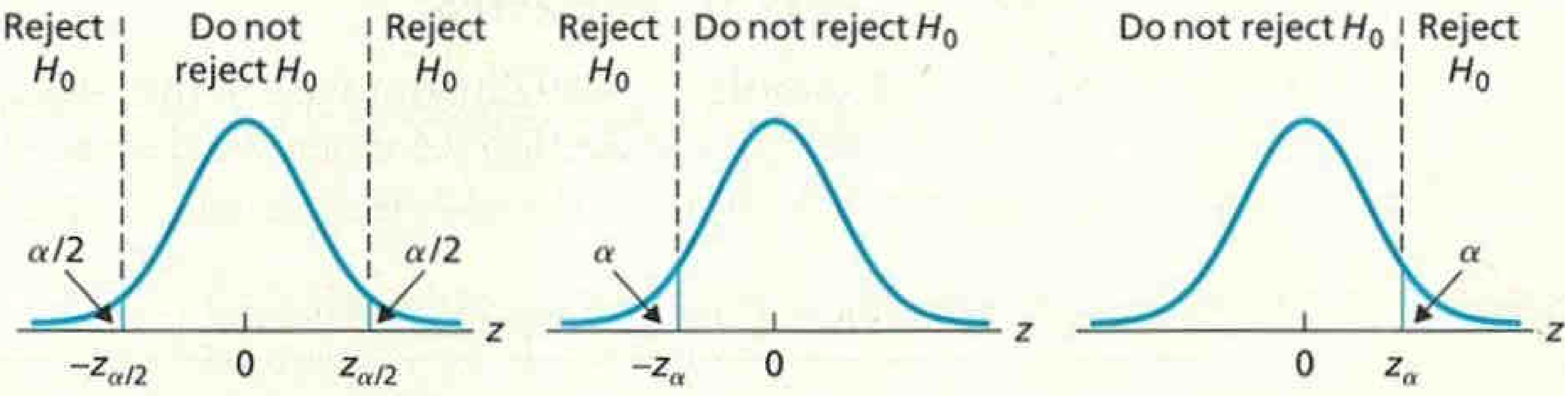
$\pm z_{\alpha/2}$		$-z_{\alpha}$		z_{α}
(Two-tailed)	or	(Left-tailed)	or	(Right-tailed)

Use Table II to find the critical value(s).

Step 4 The critical value(s) are

$$\pm z_{\alpha/2} \text{ (Two-tailed)} \quad \text{or} \quad -z_\alpha \text{ (Left-tailed)} \quad \text{or} \quad z_\alpha \text{ (Right-tailed)}$$

Use Table II to find the critical value(s).



Two-tailed — Left-tailed — Right-tailed

Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

When to Use the z-Test[2]

- For small samples—say, of size less than 15—the z-test should be used only when the variable under consideration is normally distributed or very close to being so.
- For samples of moderate size—say, between 15 and 30—the z-test can be used unless the data contain outliers or the variable under consideration is far from being normally distributed.
- For large samples—say, of size 30 or more—the z-test can be used essentially without restriction. However, if outliers are present and their removal is not justified, the effect of the outliers on the hypothesis test should be examined; that is, you should perform the hypothesis test, once with the outliers and once without them. If the conclusion remains the same either way, you may be content to take that as your conclusion and close the investigation. But if the conclusion is affected, you probably should make the more conservative conclusion, use a different procedure, or take another sample.
- If outliers are present but their removal is justified and results in a data set for which the z-test is appropriate (as previously stated), the procedure can be used.

Example 9.9 *The One-Sample z-Test*

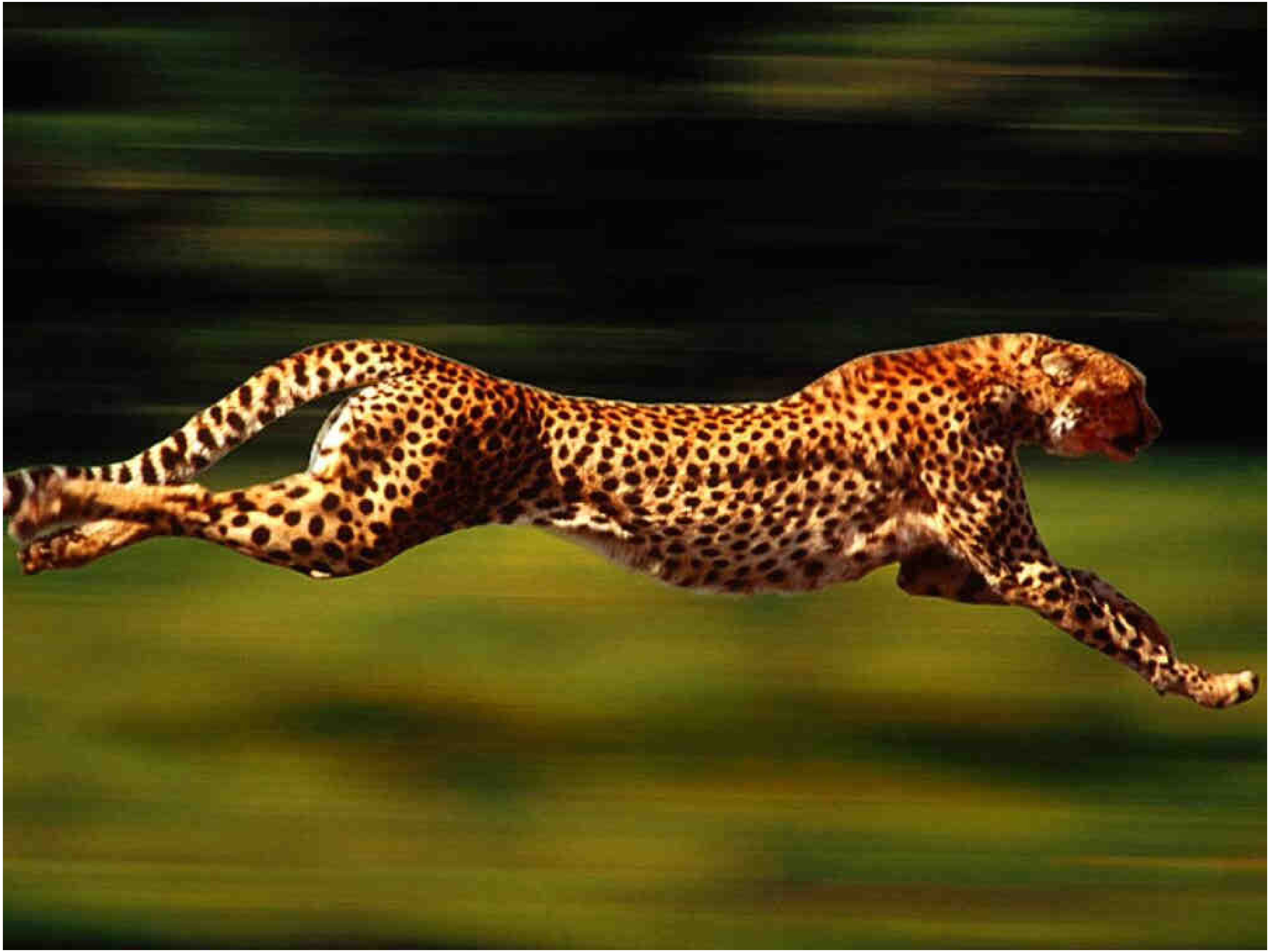
Clocking the Cheetah The Cheetah (*Acinonyx jubatus*) is the fastest land mammal on earth and is highly specialized to run down prey. According to the Cheetah Conservation of Southern Africa *Trade Environment Database*, the cheetah often exceeds speeds of 60 miles per hour (mph) and has been clocked at speeds of more than 70 mph.

One common estimate of mean top speed for cheetahs is 60 mph. Table 9.7 gives the top speeds, in miles per hour, over a quarter mile for a sample of 35 cheetahs. Do the data provide sufficient evidence to conclude that the mean top speed of all cheetahs differs from 60 mph? Assume that the population standard deviation of top speeds is 3.2 mph.

Η **τυπική απόκλιση** του **πληθυσμού** είναι γνωστή!

TABLE 9.7
Top speeds, in miles per hour, for a sample of 35 cheetahs

57.3	57.5	59.0	56.5	61.3	57.6	59.2
65.0	60.1	59.7	62.6	52.6	60.7	62.3
65.2	54.8	55.4	55.5	57.8	58.7	57.8
60.9	75.3	60.6	58.1	55.9	61.6	59.6
59.8	63.4	54.7	60.2	52.4	58.3	66.0



Statistix 9.0 cheetahs_12May09, 12-May-09, 01:38:07 PM

Descriptive Statistics

	SPEED_MPH
N	35
Lo 95% CI	58.058
Mean	59.526
Up 95% CI	60.994
SD	4.2739
SE Mean	0.7224
Minimum	52.400
Maximum	75.300

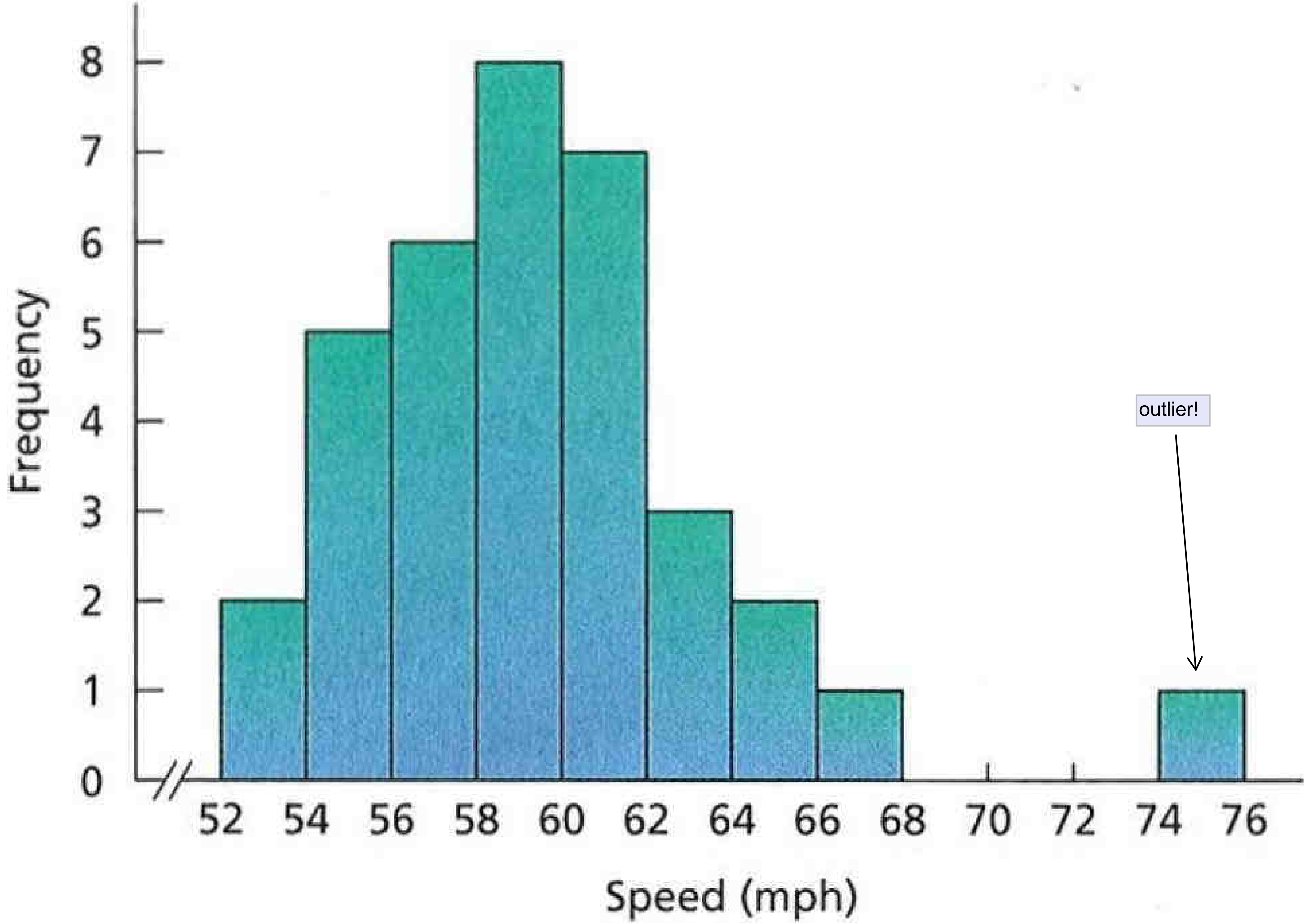
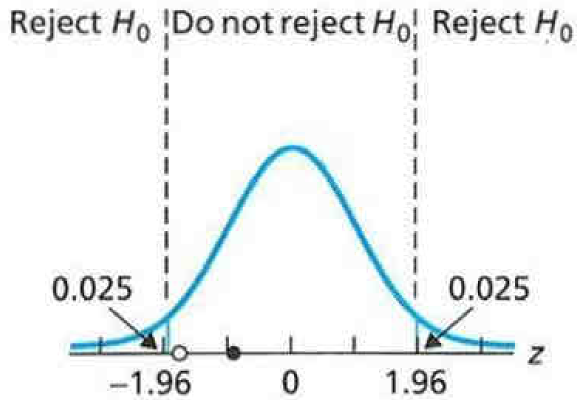

Frequency
8
7
6
5
4
3
2
1
0
52
54
56
58
60
62
64
66
68
70
72
74
76
Speed (mph)
outlier!

FIGURE 9.12
Criterion for deciding whether to reject the null hypothesis



Step 4 The critical values for a two-tailed test are $\pm z_{\alpha/2}$.

As $\alpha = 0.05$, we obtain from Table II (or Table 9.4 or Table IV) the critical values of $\pm z_{0.05/2} = \pm z_{0.025} = \pm 1.96$, as shown in Fig. 9.12.

Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

From Step 3, the value of the test statistic is $z = -0.88$. This value does not fall in the rejection region shown in Fig. 9.12. Hence we do not reject H_0. The test results are not statistically significant at the 5% level.

Step 3 **Compute the value of the test statistic**

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.$$

We have $\mu_0 = 60$, $\sigma = 3.2$, and $n = 35$. From the data in Table 9.7, we find that $\bar{x} = 59.526$. Thus the value of the test statistic is

$$z = \frac{59.526 - 60}{3.2/\sqrt{35}} = -0.88.$$

This value of z is marked with a solid dot in Fig. 9.12.

Step 6 Interpret the results of the hypothesis test.

At the 5% significance level, the (complete) data do not provide sufficient evidence to conclude that the mean top speed of all cheetahs differs from 60 mph.

We have now completed the hypothesis test, using all 35 top speeds in Table 9.7. However, recall that the top speed of 75.3 mph is an outlier. For this problem, although we don't actually know whether removing this outlier is justified (a common situation), we can still remove it from the sample data and assess the effect on the hypothesis test.

Doing so, we find that the value of the test statistic for the abridged data is $z = -1.71$, which we have marked with a hollow dot in Fig. 9.12. This value still lies in the nonrejection region, although it is much closer to the rejection region than the value of the test statistic for the unabridged data, $z = -0.88$.

In this case, removing the outlier therefore does not affect the conclusion of the hypothesis test. We can probably accept that the mean top speed of all cheetahs is roughly 60 mph.