

One-Sample Tests of Hypothesis

Chapter 10

Learning Objectives

- LO10-1 Explain the process of testing a hypothesis
- LO10-2 Apply the six-step procedure for testing a hypothesis
- LO10-3 Distinguish between a one-tailed and a two-tailed test of hypothesis
- LO10-4 Conduct a test of a hypothesis about a population mean
- LO10-5 Compute and interpret a p-value
- LO10-6 Use a t-statistic to test a hypothesis
- LO10-7 Compute the probability of a Type II error

Hypothesis Testing

- ▶ Hypothesis testing begins with a hypothesis statement about a population parameter

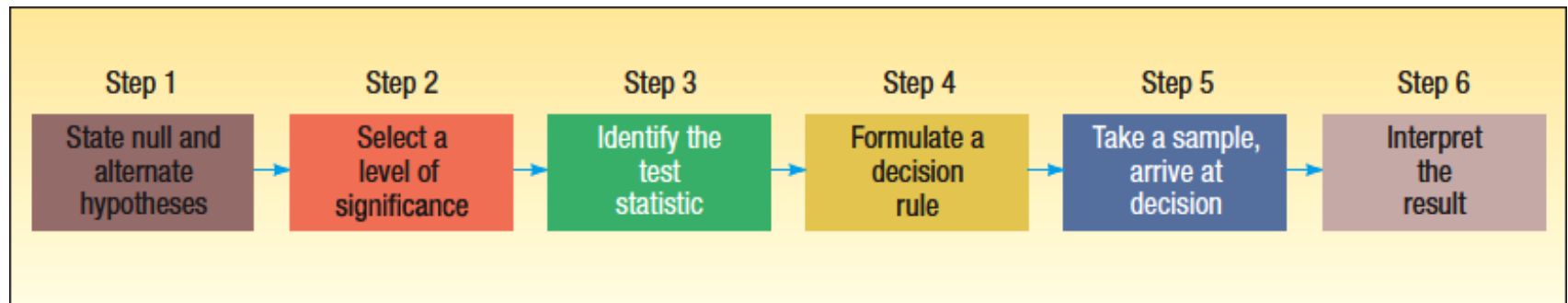
HYPOTHESIS A statement about a population parameter subject to verification

- ▶ Examples
- ▶ The mean speed of automobiles passing milepost 150 on the West Virginia Turnpike is 68 mph
- ▶ The mean cost to remodel a kitchen is \$20,000

Hypothesis Testing

- ▶ The objective of hypothesis testing is to verify the validity of a statement about a population parameter

HYPOTHESIS TESTING A procedure based on sample evidence and probability theory to determine whether the hypothesis is a reasonable statement.



Step 1 of the Six-Step Process

- ▶ State the null hypothesis (H_0) and the alternate hypothesis (H_1)

NULL HYPOTHESIS A statement about the value of a population parameter developed for the purpose of testing numerical evidence.

- ▶ The null hypothesis always includes the equal sign
 - ▶ For example; $=$, $\geq$, or $\leq$ will be used in H_0

ALTERNATE HYPOTHESIS A statement that is accepted if the sample data provide sufficient evidence that the null hypothesis is false.

- ▶ The alternate hypothesis never includes the equal sign
 - ▶ For example; $\neq$, $<$, or $>$ is used in H_1

Step 2 of the Process

- ▶ Next, you select the level of significance, α

LEVEL OF SIGNIFICANCE The probability of rejecting the null hypothesis when it is true.

- ▶ Sometimes called the level of risk
- ▶ Can be any value between 0 and 1
- ▶ Traditionally,
 - ▶ .05 is used for consumer research projects
 - ▶ .01 for quality assurance
 - ▶ .10 for political polling

Possible Error in Hypothesis Testing

- ▶ Since the researcher cannot study every item or individual in the population, error is possible

TYPE I ERROR Rejecting the null hypothesis, H_0 , when it is true.

- ▶ Type I error is designated with the Greek letter alpha, α

TYPE II ERROR Not rejecting the null hypothesis when it is false.

- ▶ Type II error is designated with the Greek letter beta, β

Null Hypothesis	Researcher	
	Does Not Reject H_0	Rejects H_0
H_0 is true	Correct decision	Type I error
H_0 is false	Type II error	Correct decision

Step 3 of the Process

- ▶ Then, select the test statistic

TEST STATISTIC A value, determined from sample information, used to determine whether to reject the null hypothesis.

- ▶ In hypothesis testing for the mean, μ , when σ is known, the test statistic z is computed with the following formula

TESTING A MEAN, σ KNOWN

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

(10-1)

- ▶ We can determine whether the distance between $\bar{x}$ and μ is statistically significant by finding the number of standard deviations $\bar{x}$ is from μ

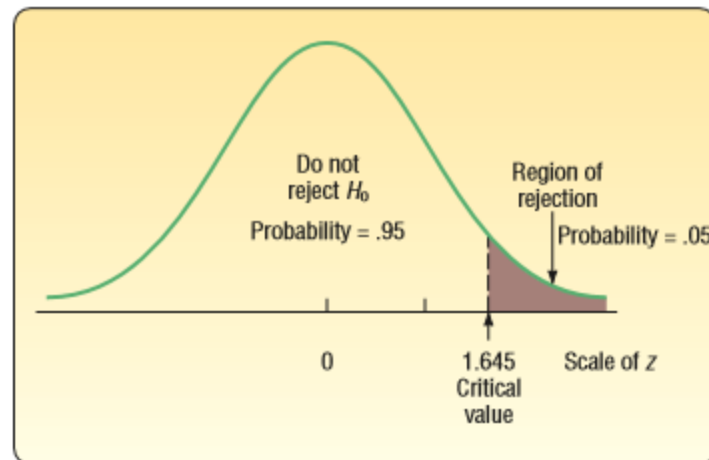
Step 4 of the Process

- ▶ Formulate the decision rule
- ▶ The decision rule is a statement of specific conditions under which the null hypothesis is rejected and the conditions under which it is not rejected
- ▶ The region or area of rejection defines the location of all the values that are either so large or so small that their probability of occurrence under a true null hypothesis is remote

CRITICAL VALUE The dividing point between the region where the null hypothesis is rejected and the region where it is not rejected.

Critical Value

- ▶ The sampling distribution of the statistic z follows the normal distribution
- ▶ Here, an α of .05 is used in a one-tailed test
- ▶ The value 1.645 separates the regions where the null hypothesis is rejected and where it is not rejected
- ▶ The value 1.645 is the critical value



Steps 5 & 6 of the Six-Step Process

▶ Step 5 Make a decision

- ▶ First, select a sample and compute the value of the test statistic
- ▶ Compare the value of the test statistic to the critical value
- ▶ Then, make the decision regarding the null hypothesis

▶ Step 6 Interpret the results

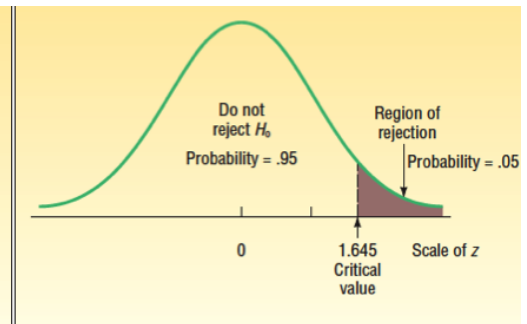
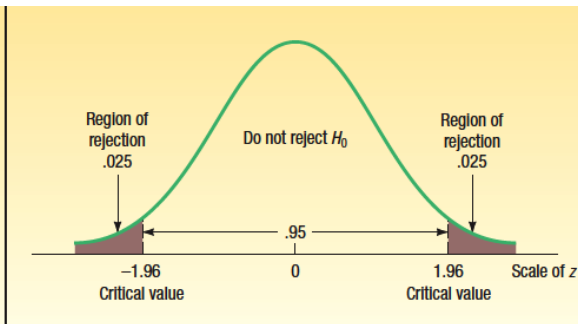
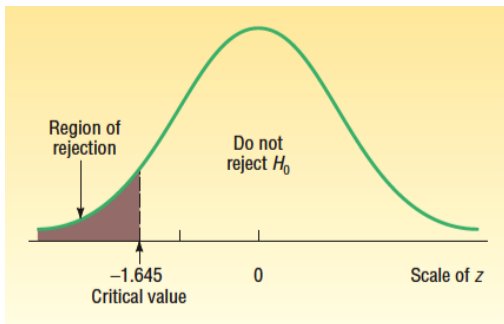
- ▶ What can we say or report based on the results of the statistical test?

One-Tailed and Two-Tailed Tests

$H_0: \geq 60,000$ miles
 $H_1: < 60,000$ miles
with an $\alpha = .05$
Left-tailed test

$H_0: = \$65,000$ per year
 $H_1: \neq \$65,000$ per year
with an $\alpha = .05$
Two-tailed test

$H_0: \leq 453$ grams
 $H_1: > 453$ grams
with an $\alpha = .05$
Right-tailed test



Note that the total area in the normal distribution is 1.0000.

Two-Tailed Test Example, σ Known

Jamestown Steel Company manufactures and assembles desks and other office equipment at several plants in New York State. At the Fredonia plant, the weekly production of the Model A325 desk follows a normal distribution with a mean of 200 and a standard deviation of 16. New production methods have been introduced and the vice president of manufacturing would like to investigate whether there has been a change in weekly production of the Model A325. Is the mean number of desks produced different from 200 at the .01 significance level?

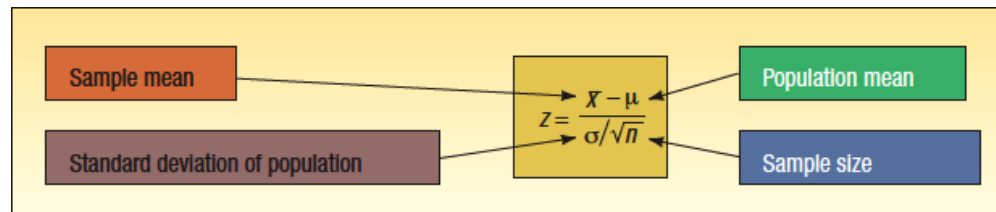
Step 1: State the null hypothesis and alternate hypothesis.

$$H_0: \mu = 200 \text{ desks}$$

$$H_1: \mu \neq 200 \text{ desks}$$

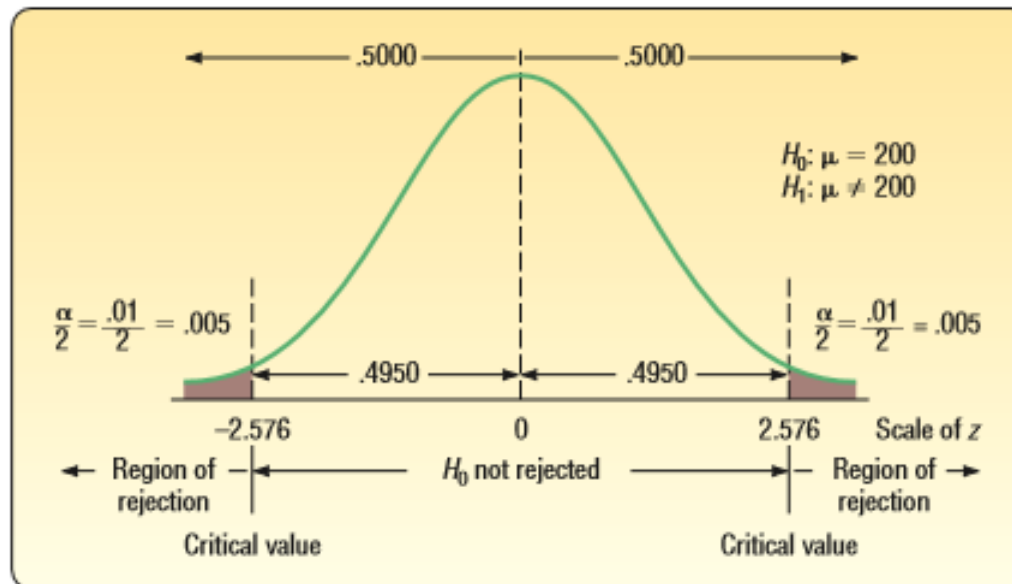
Step 2: Select the level of significance. Here $\alpha = .01$

Step 3: Select the test statistic. In this example, we'll use z



Two-Tailed Test Example, σ Known

Step 4: Formulate the decision rule by first determining the critical values of z .



Decision Rule: If the computed value of z is not between -2.576 and 2.576 , reject the null hypothesis. If z falls between -2.576 and 2.576 , do not reject the null hypothesis.

Two-Tailed Test Example, σ Known

Step 5: Take sample, compute the test statistic, make decision.

The mean number of desks produced last year (50 weeks because the plant was shut down 2 weeks for vacation) is 203.5. The standard deviation of the population is 16 desks per week. Compute z with formula 10-1.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{203.5 - 200}{16/\sqrt{50}} = 1.547$$

Decision: Because 1.547 does not fall in the rejection region, we decide not to reject H_0 .

Step 6: Interpret the result.

We did not reject the null hypothesis, so we have failed to show that the population mean has changed from 200 per week.

One-Tailed Test

Suppose instead of wanting to know if there had been a change in the mean number of desks assembled, the vice president wanted to know if there had been an increase in the number of units assembled. Can we conclude, because of the improved production methods, that the mean number of desks assembled in the last 50 weeks was more than 200? Use $\alpha = .01$.

Before:

A two-tailed test

$H_0: \mu = 200$ desks

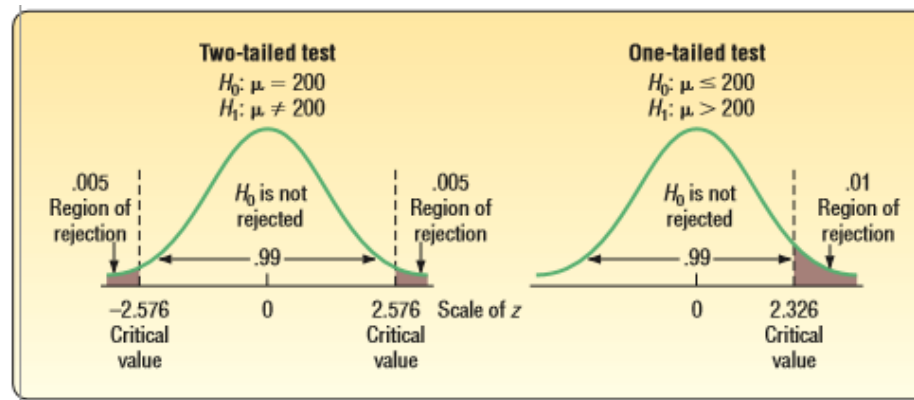
$H_1: \mu \neq 200$ desks

Now:

A one-tailed test

$H_0: \mu \leq 200$ desks

$H_1: \mu > 200$ desks



The p-Value in Hypothesis Testing

p-VALUE The probability of observing a sample value as extreme as, or more extreme than the value observed, given that the null hypothesis is true.

- ▶ Compare the p-value with the level of significance, α
 - ▶ If the p-value is smaller than the significance level, reject H_0
 - ▶ If the p-value is larger than α , H_0 is not rejected
- ▶ A p-value not only results in a decision about H_0 , but gives additional insight about the strength of that decision

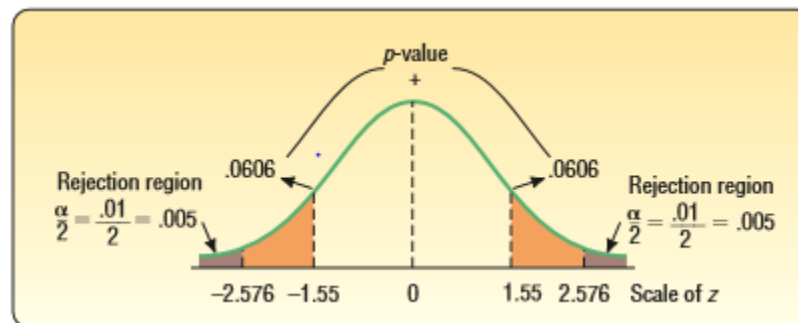
INTERPRETING THE WEIGHT OF EVIDENCE AGAINST H_0

If the p-value is less than

- (a) .10, we have *some* evidence that H_0 is not true.
- (b) .05, we have *strong* evidence that H_0 is not true.
- (c) .01, we have *very strong* evidence that H_0 is not true.
- (d) .001, we have *extremely strong* evidence that H_0 is not true.

Finding a p-Value

- ▶ In the previous example about desk production, the computed z was 1.547 and H_0 was not rejected
- ▶ Round the computed z -value to two decimal places, 1.55
- ▶ Using the z -table, find the probability of finding a z -value of 1.55 or more by $.5000 - .4394 = .0606$
- ▶ Since this is a two-tailed test $2(.0606) = .1212$
- ▶ In this chart, we can easily compare the p -value with the level of significance



Exercise 1

Answer the questions: (a) Is this a one- or two-tailed test? (b) What is the decision rule? (c) What is the value of the test statistic? (d) What is your decision regarding H_0 ? (e) What is the p -value? Interpret it.

1. A sample of 36 observations is selected from a normal population. The sample mean is 49, and the population standard deviation is 5. Conduct the following test of hypothesis using the .05 significance level.

$$\begin{array}{l} H_0: \mu = 50 \\ H_1: \mu \neq 50 \end{array}$$

a. Two-tailed, because the alternate hypothesis does not indicate a direction.

b. Reject H_0 when z does not fall in the region from -1.96 and 1.96

c. -1.2 , found by $z = \frac{49 - 50}{(5 / \sqrt{36})}$

d. Fail to reject H_0

e. $p = 0.2302$, found by $2(0.5000 - 0.3849)$. There is a 23.02% chance of finding a z value this large by “sampling error” when H_0 is true

Exercise 2

Answer the questions: (a) Is this a one- or two-tailed test? (b) What is the decision rule? (c) What is the value of the test statistic? (d) What is your decision regarding H_0 ? (e) What is the p -value? Interpret it.

→ A sample of 36 observations is selected from a normal population. The sample mean is 12, and the population standard deviation is 3. Conduct the following test of hypothesis using the .02 significance level.

$$\begin{array}{l} H_0: \mu \leq 10 \\ H_1: \mu > 10 \end{array}$$

a. One-tailed, because the alternate hypothesis indicates a greater than direction

b. Reject H_0 when $z > 2.05$

c. 4 found by $z = \frac{12 - 10}{(3 / \sqrt{36})}$

d. Reject H_0 and conclude that $\mu > 10$

e. The p -value is close to 0. So there is very little chance H_0 is true

Exercise 3

The waiting time for customers at MacBurger Restaurants follows a normal distribution with a mean of 3 minutes and a standard deviation of 1 minute. At the Warren Road MacBurger, the quality-assurance department sampled 50 customers and found that the mean waiting time was 2.75 minutes. At the .05 significance level, can we conclude that the mean waiting time is less than 3 minutes?

- a) State the null hypothesis and the alternate hypothesis.
- b) State the decision rule.
- c) Compute the value of the test statistic.
- d) What is your decision regarding H_0 ?
- e) What is the p -value? Interpret it.

a. $H_0: \mu \geq 3$ $H_1: \mu < 3$

b. Reject H_0 if $z < -1.65$

c. -1.77 , found by $z = \frac{2.75 - 3.0}{(1/\sqrt{50})}$

d. Reject H_0

e. $p = 0.0384$, found by $(0.5000 - 0.4616)$. We conclude that the mean waiting time is less than three minutes. When H_0 is true, the probability of obtaining a value smaller than -1.77 is 0.0384.

Hypothesis Testing, σ Unknown

- ▶ When testing a hypothesis about a population mean

TESTING A MEAN, σ UNKNOWN

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

[10-2]

- ▶ The major characteristics of the t distribution are
 - ▶ It is a continuous distribution
 - ▶ It is bell-shaped and symmetrical
 - ▶ There is a family of t distributions, depending on the number of degrees of freedom
 - ▶ It is flatter, or more spread out, than the standard normal distribution

Hypothesis Testing, σ Unknown Example

The McFarland Insurance Company Claims Department reports the mean cost to process a claim is \$60. An industry comparison showed this amount to be larger than most other insurance companies, so the company instituted cost-cutting measures. To evaluate the cost-cutting measures, a random sample was taken of 26 claims processed last month and the cost to process each claim was recorded (see below).

At the .01 significance level, is it reasonable to conclude the mean cost to process a claim is now less than \$60?

\$45	\$49	\$62	\$40	\$43	\$61
48	53	67	63	78	64
48	54	51	56	63	69
58	51	58	59	56	57
38	76				

Step 1: State the null hypothesis and the alternate hypothesis

$$H_0: \mu \geq \$60$$

$$H_1: \mu < \$60$$

Hypothesis Testing, σ Unknown Example

Step 2: Select the level of significance; we will use .01

Step 3: Select the test statistic; we will use t

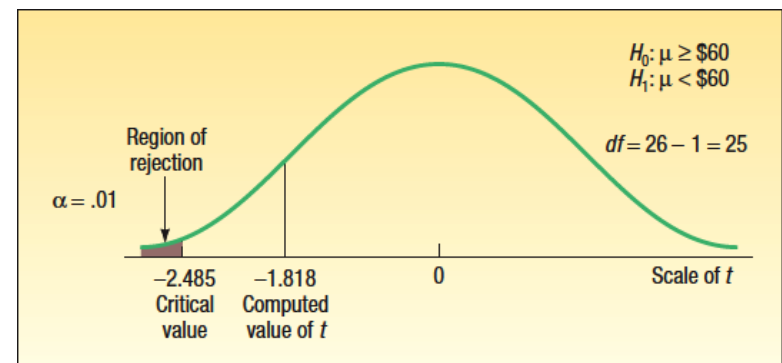
Step 4: Formulate the decision rule; reject H_0 if t is less than -2.485

Step 5: Take sample, make decision; Do not reject H_0

Step 6: Interpret the result; The test results do not allow the claims manager to conclude the cost-cutting measures have been effective.

Confidence Intervals						
	80%	90%	95%	98%	99%	99.9%
df	Level of Significance for One-Tailed Test, α					
	0.10	0.05	0.025	0.01	0.005	0.0005
	Level of Significance for Two-Tailed Test, α					
	0.20	0.10	0.05	0.02	0.01	0.001
...	...	...	...	...	...	...
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.768
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{\$56.423 - \$60}{\$10.41/\sqrt{26}} = -1.82$$



Exercise (1 / 2)

Most air travelers now use e-tickets. Electronic ticketing allows passengers to not worry about a paper ticket, and it costs the airline companies less to handle than paper ticketing. However, in recent times the airlines have received complaints from passengers regarding their e-tickets, particularly when connecting flights and a change of airlines were involved. To investigate the problem, an independent watchdog agency contacted a random sample of 20 airports and collected information on the number of complaints the airport had with e-tickets for the month of March. The information is reported below.

14	14	16	12	12	14	13	16	15	14
12	15	15	14	13	13	12	13	10	13

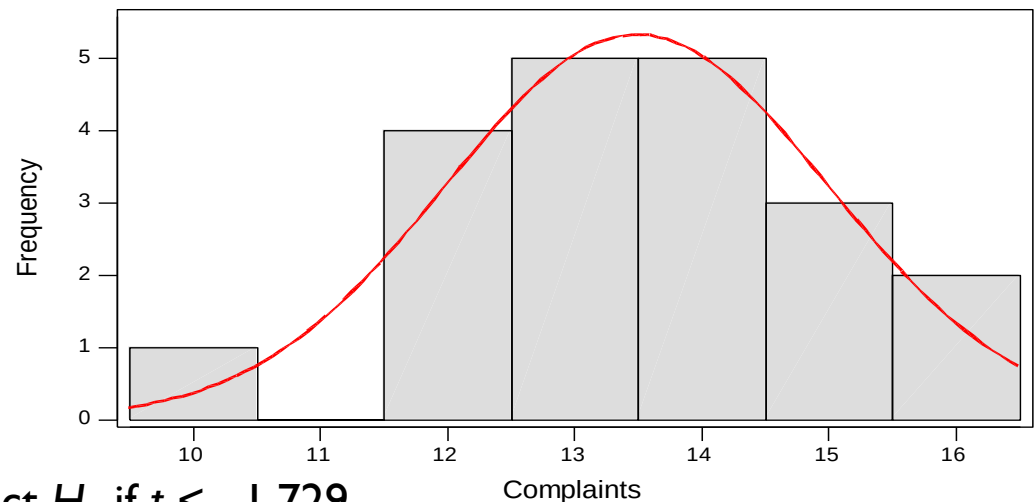
At the .05 significance level, can the watchdog agency conclude the mean number of complaints per airport is less than 15 per month?

- What assumption is necessary before conducting a test of hypothesis?
- Plot the number of complaints per airport in a frequency distribution or a dot plot. Is it reasonable to conclude that the population follows a normal distribution?
- Conduct a test of hypothesis and interpret the results.

Exercise (2/2)

14	14	16	12	12	14	13	16	15	14
12	15	15	14	13	13	12	13	10	13

Histogram of Complaints, with Normal Curve



- a. The population of complaints follows a normal probability distribution.
- b. The assumption of normality appears reasonable.

c. $H_0 : \mu \geq 15$ $H_1 : \mu < 15$ Reject H_0 if $t < -1.729$

$$t = \frac{(13.5 - 15)}{(1.504 / \sqrt{20})} = -4.46$$

Reject the null hypothesis. The mean number of complaints is less than 15

Type I and Type II Errors

- ▶ Type I error occurs when a true null hypothesis is rejected
 - ▶ The probability of making a Type I error is equal to the level of significance, α
- ▶ A Type II error occurs when a false null hypothesis is not rejected
 - ▶ The probability of making a Type II error is designated, β
 - ▶ The likelihood of a Type II error must be calculated comparing the hypothesized distribution to an alternate distribution based on sample results and can be calculated with this formula

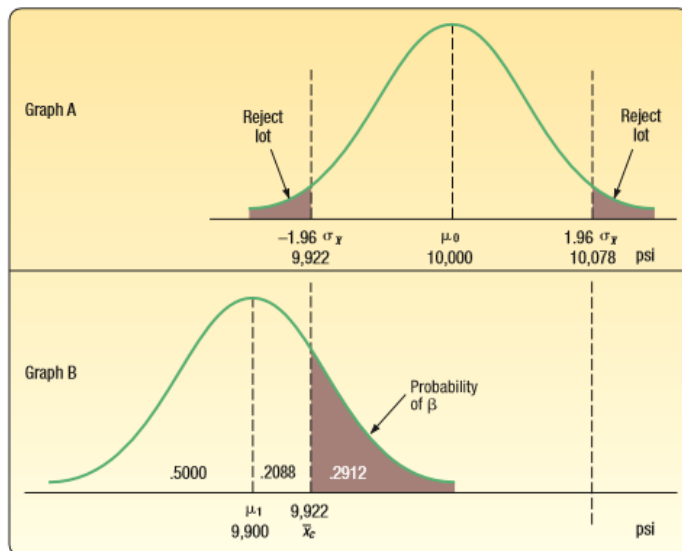
TYPE II ERROR

$$z = \frac{\bar{X}_c - \mu_1}{\sigma/\sqrt{n}}$$

(10-3)

Type II Error Example

Western Wire Products purchases steel bars to make cotter pins. Past experience indicates that the mean tensile strength of all incoming shipments is 10,000 psi and that the standard deviation is 400 psi. To monitor the quality of the cotter pins, samples of 100 pins are randomly selected and tested for their strength. Using a 0.05 significance level, accept the shipment if the sample mean strength falls between the critical values 9.922 psi and 10.078 psi. If the sample mean does not fall between the critical values, we conclude the shipment does not meet the quality standard.



The sample mean, 9.900 psi, is not within the specified range. To calculate the probability of a Type II error, assume the sample mean is the true mean (see graph B). Determine the probability of the sample mean falling between 9.900 and 9.922. Then subtract this probability from .5000 to arrive at the probability of making a Type II error, .2912