LECTURE: 4

Hypothesis Test

Professor Dr.Heider Al-Heidery

M.B.Ch.B., M.Sc., Ph.D.

Definitions

In statistics, a hypothesis is a claim or statement about a property of a population.

A hypothesis test is a standard procedure for testing a claim about a property of a population.

Main Objectives

We will study hypothesis testing for

1. **population proportion** p
2. **population mean** μ
3. **population standard deviation** σ

Example

Claim: the XSORT method of gender selection increases the likelihood of having a baby girl.

This is a claim about proportion (of girls)

To test this claim 14 couples (volunteers) were subject to XSORT treatment.

If 6 or 7 or 8 have girls, the method probably does not increase the likelihood of a girl.

If 13 or 14 couples have girls, the method is probably increases the likelihood of a girl.

Rare Event Rule for Inferential Statistics

If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.

Components of a Formal Hypothesis Test

Null Hypothesis: H_0

- 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.
- We test the null hypothesis directly.
- Either reject H_0 or fail to reject H_0

(in other words, accept H_0).

Alternative Hypothesis: H_1

- The alternative hypothesis (denoted by H_1) is the statement that the parameter has a value that somehow differs from the null hypothesis.
- The symbolic form of the alternative hypothesis must use one of these symbols: $\neq$, $<$, $>$.
(not equal, less than, greater than)

Example 1

Claim: the XSORT method of gender selection increases the likelihood of having a baby girl.

We express this claim in symbolic form: $p>0.5$
(here p denotes the proportion of baby girls)

Null hypothesis must say "equal to", so

$$H_0 : p=0.5$$

Alternative hypothesis must express difference:

$$H_1 : p>0.5$$

Original claim is now the alternative hypothesis

Example 1 (continued)

We always test the null hypothesis.

If we reject the null hypothesis, then the original clam is accepted.

Final conclusion would be: XSORT method increases the likelihood of having a baby girl.

If we fail to reject the null hypothesis, then the original clam is rejected.

Final conclusion would be: XSORT method does not increase the likelihood of having a baby girl.

Example 2

<u>Claim</u>: for couples using the XSORT method the likelihood of having a baby girl is 50%

Express this claim in symbolic form: p=0.5
(again p denotes the proportion of baby girls)

Null hypothesis must say "equal to", so

$$H_0 : p=0.5$$

Alternative hypothesis must express difference:

$$H_1 : p\neq 0.5$$

Original claim is now the null hypothesis

Example 2 (continued)

If we reject the null hypothesis, then the original clam is rejected.

Final conclusion would be: for couples using the XSORT, the likelihood of having a baby girl is not 0.5

If we fail to reject the null hypothesis, then the original clam is accepted.

Final conclusion would be: for couples using the XSORT the likelihood of having a baby girl is indeed equal to 0.5

Example 3

Claim: for couples using the XSORT method the likelihood of having a baby girl is **at least 0.5**

Express this claim in symbolic form: $p \geq 0.5$
(again *p* denotes the proportion of baby girls)

Null hypothesis must say "equal to", so

$H_0 : p = 0.5$ **(this agrees with the claim!)**

Alternative hypothesis must express difference:

$H_1 : p < 0.5$

Original claim is now the null hypothesis

Example 3 (continued)

If we reject the null hypothesis, then the original clam is rejected.

Final conclusion would be: for couples using the XSORT, the likelihood of having a baby girl is less 0.5

If we fail to reject the null hypothesis, then the original clam is accepted.

Final conclusion would be: for couples using the XSORT the likelihood of having a baby girl is indeed at least 0.5

General rules:

- **If the null hypothesis is <u>rejected</u>, the alternative hypothesis is <u>accepted</u>.**
- **If the null hypothesis is <u>accepted</u>, the alternative hypothesis is <u>rejected</u>.**
- **Acceptance or rejection of the null hypothesis is an initial conclusion.**
- **Always state the final conclusion expressed in terms of the original claim, not in terms of the null hypothesis or the alternative hypothesis.**

Type I Error

- A **Type I error** is the mistake of rejecting the null hypothesis when it is actually true.

- The symbol α (alpha) is used to represent the probability of a type I error.

Type II Error

- A **Type II error** is the mistake of accepting the null hypothesis when it is actually false.
- The symbol β (beta) is used to represent the probability of a type II error.

Type I and Type II Errors

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	**Type I error** (rejecting a true null hypothesis) $P(\text{type I error}) = \alpha$	Correct decision
	We fail to reject the null hypothesis	Correct decision	**Type II error** (failing to reject a false null hypothesis) $P(\text{type II error}) = \beta$

Example

Claim:

a new medicine has a greater success rate, $p>p_0$, than the old (existing) one.

Null hypothesis: $H_0 : p=p_0$

Alternative hypothesis: $H_1 : p>p_0$

(agrees with the original claim)

Example (continued)

Type I error: the null hypothesis is true, but we reject it => we accept the claim, hence we adopt the new (<u>inefficient, potentially harmful</u>) medicine.

This is a critical error, should be avoided!

Type II error: the alternative hypothesis is true, but we reject it => we reject the claim, hence we decline the new medicine and continue using the old one (no harm…).

Significance Level

The probability of the type I error (denoted by α) is also called the **significance level** of the test.

It characterizes the chances that the test fails (i.e., type I error occurs)

It must be a small number. Typical values used in practice: α = 0.1, 0.05, or 0.01 (in percents, 10%, 5%, or 1%).

Testing hypothesis
Step 1: compute Test Statistic

The test statistic is a value used in making a decision about the null hypothesis.

The test statistic is computed by a specific formula depending on the type of the test.

Section 8-3

Testing a Claim About a Proportion

Notation

n = number of trials

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

p = population proportion (must be specified in the null hypothesis)

$q = 1 - p$

Requirements for Testing Claims About a Population Proportion *p*

1) The sample observations are a simple random sample.

2) The conditions for a binomial distribution are satisfied.

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 $\sigma = \sqrt{npq}$. Note: *p* is the assumed proportion not the sample proportion.

Test Statistic for Testing a Claim About a Proportion

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

Note: p **is the value specified in the null hypothesis;** $q = 1 - p$

Example 1 again:

Claim: the XSORT method of gender selection increases the likelihood of having a baby girl.

Null hypothesis: $H_0 : p = 0.5$
Alternative hypothesis: $H_1 : p > 0.5$

Suppose 14 couples treated by XSORT gave birth to 13 girls and 1 boy.

Test the claim at a 5% significance level

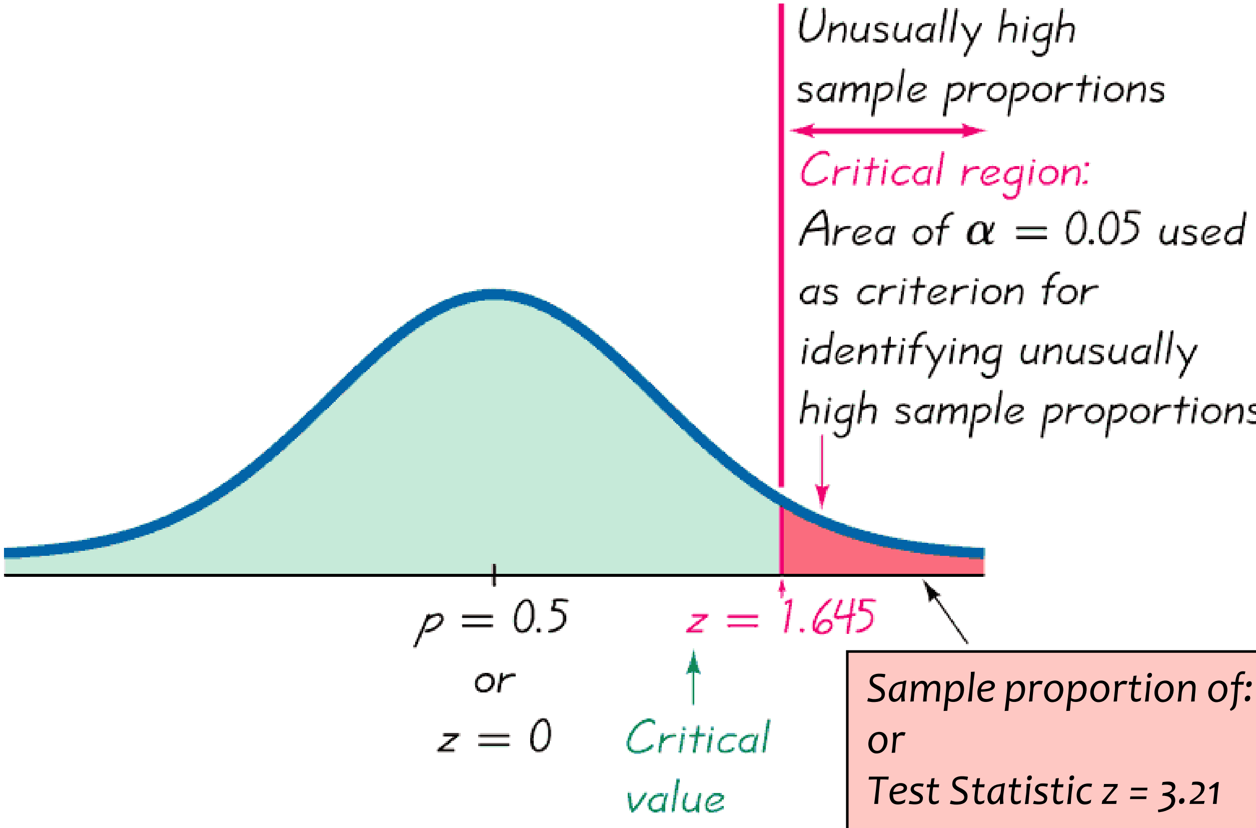
Compute the test statistic:

$$\hat{p} = 13/14 = 0.929$$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.929 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{14}}} = 3.21$$

Draw the diagram (the normal curve)

On the diagram, mark a region of extreme values that agree with the alternative hypothesis:



Sample proportion of: $\hat{p} = 0.929$
or
Test Statistic z = 3.21

Critical Region

The critical region (or rejection region) is the set of all values of the test statistic that cause us to reject the null hypothesis.

For example, see the red-shaded region in the previous figure.

Critical Value

A critical value is a value that 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.

See the previous figure where the critical value is $z = 1.645$. It corresponds to a significance level of $\alpha = 0.05$.

Significance Level

The significance level (denoted by α) is the probability that the test statistic will fall in the critical region (when the null hypothesis is actually true).

Conclusion of the test

Since the test statistic (z=3.21) falls in the critical region (z>1.645), we **reject the null hypothesis**.

Final conclusion: the original claim is accepted, the XSORT method of gender selection indeed increases the likelihood of having a baby girl.

Types of Hypothesis Tests: Two-tailed, Left-tailed, Right-tailed

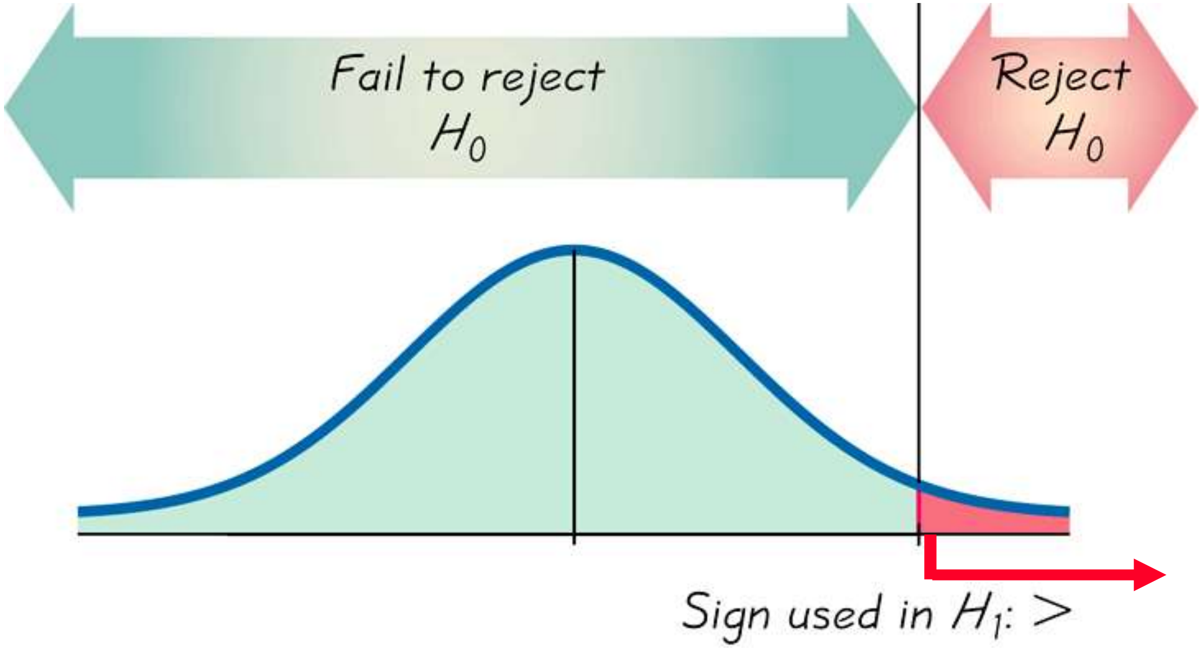
The tails in a distribution are the extreme regions where values of the test statistic agree with the alternative hypothesis

Right-tailed Test

H_0: p=0.5 α is in the right tail

H_1: p>0.5

Points Right



Critical value for a right-tailed test

A right-tailed test requires one (positive) critical value:

$$z_\alpha$$

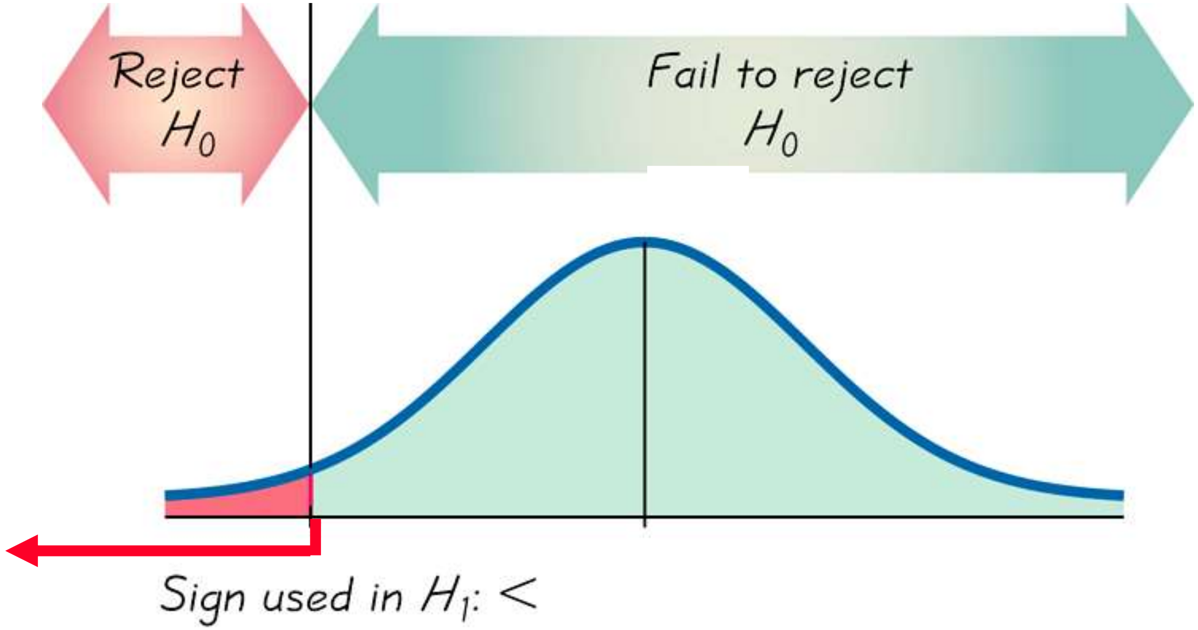
Left-tailed Test

H_0: p=0.5

H_1: p<0.5

α is in the left tail

Points Left



Critical value for a left-tailed test

A left-tailed test requires one (negative) critical value:

$$-z_{\alpha}$$

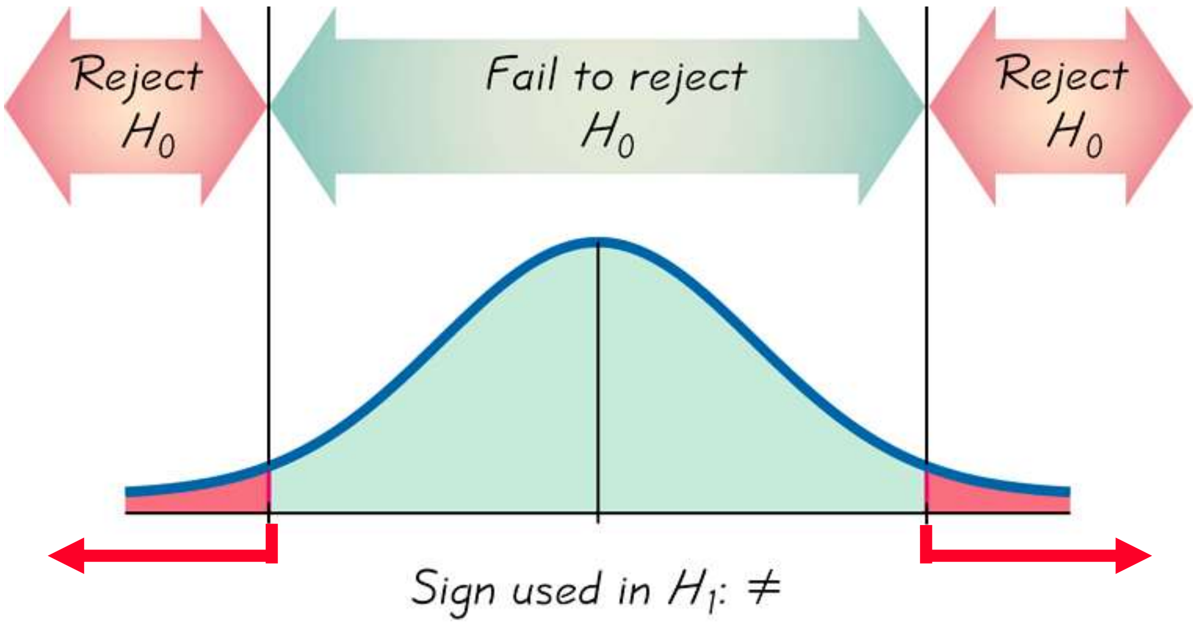
Two-tailed Test

H_0: p=0.5

H_1: p≠0.5

Means less than or greater than

α is divided equally between the two tails of the critical region



Critical values for a two-tailed test

A two-tailed test requires two critical values:

$$z_{\alpha/2} \quad \text{and} \quad -z_{\alpha/2}$$

P-Value

The *P*-value (or *p*-value or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming that the null hypothesis is true.

Example 1 (continued)

P-value is the area to the right of the test statistic z = 3.21.

We refer to Table A-2 (or use calculator) to find that the area to the right of z = 3.21 is 0.0007.

P-value = 0.0007

P-Value method:

If *P*-value $\leq \alpha$, reject H_0.

If *P*-value > α , fail to reject H_0.

If the P is low, the null must go.
If the P is high, the null will fly.

Example 1 (continued)

***P*-value = 0.0007**

It is smaller than $\alpha = 0.05$.

Hence the null hypothesis must be rejected

P-Value

Critical region in the right tail: *P*-value = area to the right of the test statistic

Critical region in the left tail: *P*-value = area to the left of the test statistic

Critical region in two tails: *P*-value = twice the area in the tail beyond the test statistic (see the following diagram)

Procedure for Finding *P*-Values

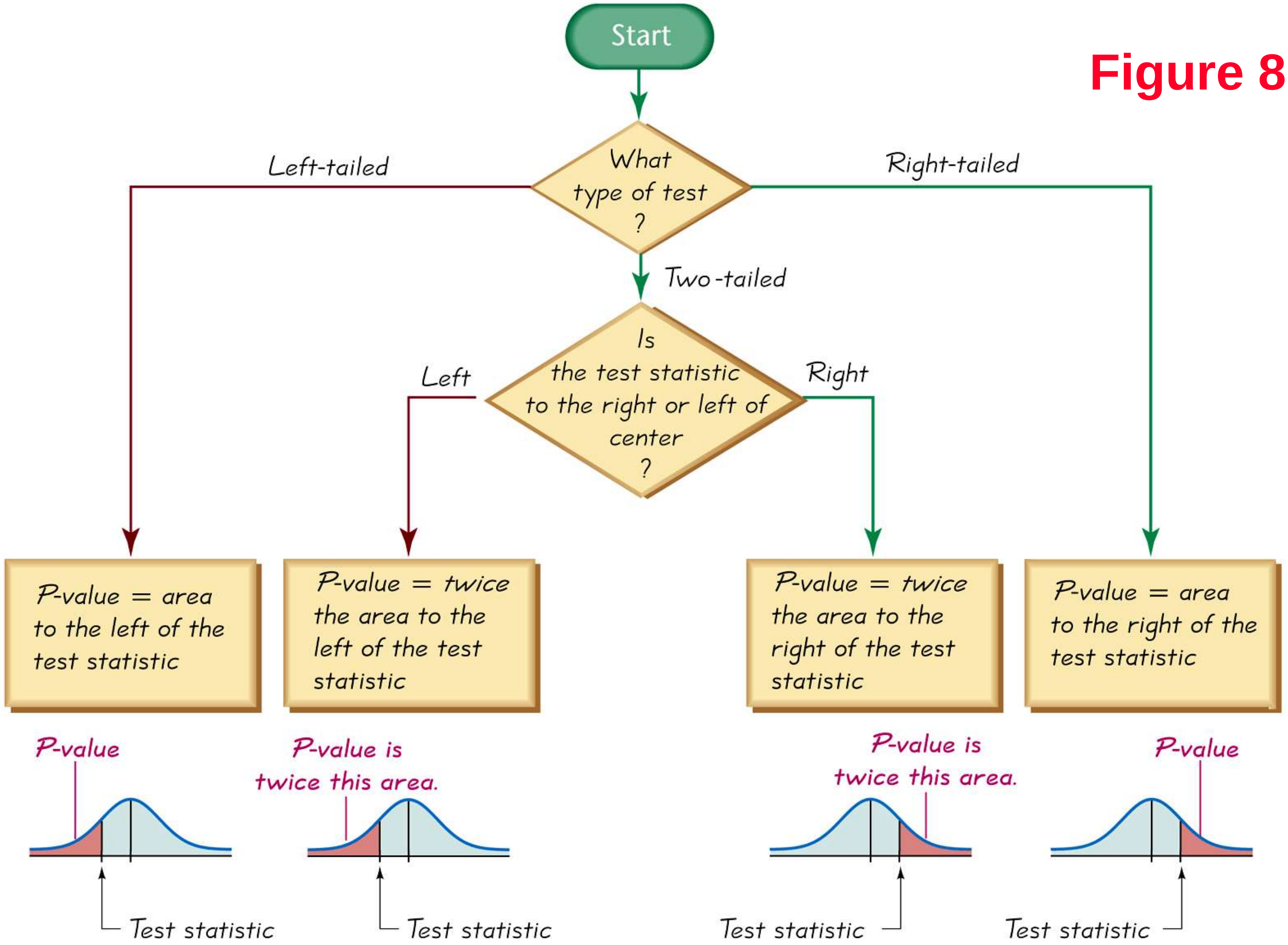


Figure 8-5

Caution

Don't confuse a *P*-value with a proportion *p*. Know this distinction:

***P*-value = probability of getting a test statistic at least as extreme as the one representing sample data**

***p* = population proportion**

Traditional method:

If the test statistic falls within the critical region, reject H_0.

If the test statistic does not fall within the critical region, fail to reject H_0 (i.e., accept H_0).

P-Value method:

If *P*-value is small ($\leq \alpha$), reject H_0.

If *P*-value is not small ($> \alpha$), accept H_0.

If the P is low, the null must go.
If the P is high, the null will fly.

Testing hypothesis by TI-83/84

- Press **STAT** and select **TESTS**
- Scroll down to **1-PropZTest** press **ENTER**
- Type in $\mathbf{p_0}$: (claimed proportion, from H_0)
- **x:** (number of successes)
- **n:** (number of trials)
- **choose $\mathbf{H_1}$:** $p \neq p_0$ $<p_0$ $>p_0$

  (two tails) (left tail) (right tail)
- Press on **Calculate**
- Read the test statistic **z=…**
- and the *P*-value **p=…**

Do we prove a claim?

- A statistical test cannot prove a hypothesis or a claim.
- Our conclusion can be only stated like this: the available evidence is not strong enough to warrant rejection of a hypothesis or a claim
(such as not enough evidence to convict a suspect).

Section 8-4

Testing a Claim About a Mean: σ Known

Notation

n = sample size

$\bar{x}$ = sample mean

μ = claimed population mean (from H_0)

σ = known value of the population standard deviation

Requirements for Testing Claims About a Population Mean (with σ Known)

1) The value of the population standard deviation σ **is known**.

2) Either or both of these conditions is satisfied: The population is **normally distributed** or $n > 30$.

Test Statistic for Testing a Claim About a Mean (with σ Known)

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

Example:

People have died in boat accidents because an obsolete estimate of the mean weight of men (166.3 lb) was used.

A random sample of n = 40 men yielded the mean $\bar{x}$ = 172.55 lb. Research from other sources suggests that the population of weights of men has a standard deviation given by σ = 26 lb.

Test the claim that men have a mean weight greater than 166.3 lb.

Example:

Requirements are satisfied: σ is known (26 lb), sample size is 40 (n > 30)

We can express claim as μ > 166.3 lb

It does not contain equality, so it is the alternative hypothesis.

H_0: μ = 166.3 lb null hypothesis
H_1: μ > 166.3 lb alternative hypothesis (and original claim)

Example:

Let us set significance level to α = 0.05

Next we calculate *z*

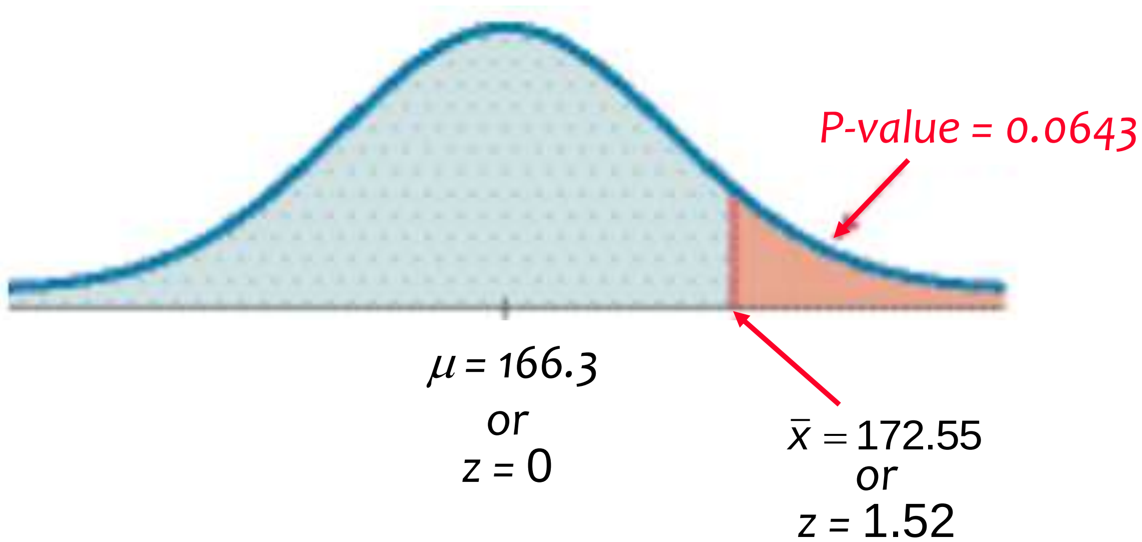
$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{172.55 - 166.3}{\frac{26}{\sqrt{40}}} = 1.52$$

It is a right-tailed test, so *P*-value is the area is to the right of *z* = 1.52;

Example:

Table A-2: area to the left of *z* = 1.52 is 0.9357, so the area to the right is 1 – 0.9357 = 0.0643.
The *P*-value is 0.0643

The *P*-value of 0.0643 is greater than the significance level of $\alpha = 0.05$, we fail to reject the null hypothesis.



Example:

The traditional method:

Use critical value $z = 1.645$ instead of finding the P-value. Since $z = 1.52$ does not fall in the critical region, again fail to reject the null hypothesis.

Testing hypothesis by TI-83/84

- Press **STAT** and select **TESTS**
- Scroll down to **Z-Test** press **ENTER**
- Choose **Data** or **Stats.** For **Stats:**
- Type in $\boldsymbol{\mu_0}$: (claimed mean, from H_0)
- **σ:** (known st. deviation)
- $\bar{x}$**:** (sample mean)
- **n:** (sample size)

choose H_1: $\mu \neq \mu_0$ $<\mu_0$ $>\mu_0$

(two tails) (left tail) (right tail)

- (continued)
- Press on **Calculate**
- Read the test statistic **z=…**
- and the *P*-value **p=…**