Introduction to Hypothesis Testing

.

Hypothesis Tests

Hypothesis test

- A process that uses sample statistics to test a claim about the value of a population parameter.
- **For example:** An automobile manufacturer advertises that its new hybrid car has a mean mileage of 50 miles per gallon. To test this claim, a sample would be taken. If the sample mean differs enough from the advertised mean, you can decide that the advertisement is wrong.

.

Hypothesis Tests

Statistical hypothesis

- A statement, or claim, about a population parameter.
- Carefully state a pair of hypotheses
  - one that represents the claim
  - the other, its complement
- When one of these hypotheses is false, the other must be true.

.

Stating a Hypothesis

Null hypothesis

- A statistical hypothesis that contains a statement of equality such as ≤, =, or ≥.
- Denoted $\boldsymbol{H_0}$ read "H subzero" or "H naught."

Alternative hypothesis

- A statement of inequality such as >, ≠, or <.
- Must be true if H_0 is false.
- Denoted $\boldsymbol{H_a}$ read "H sub-a."

.

Stating a Hypothesis

- To write the null and alternative hypotheses, translate the claim made about the population parameter from a verbal statement to a mathematical statement.
- Then write its complement.

$H_0\colon \mu \leq k$	$H_0\colon \mu \geq k$	$H_0\colon \mu = k$
$H_a\colon \mu > k$	$H_a\colon \mu < k$	$H_a\colon \mu \neq k$

- Regardless of which pair of hypotheses you use, you always assume $\mu = k$ and examine the sampling distribution on the basis of this assumption.

Example: Stating the Null and Alternative Hypotheses

Write the claim as a mathematical sentence. State the null and alternative hypotheses and identify which represents the claim.

1. A school publicizes that the proportion of its students who are involved in at least one extracurricular activity is 61%.

Solution:

H_0: $p = 0.61$ ← Equality condition (Claim)

H_a: $p \neq 0.61$ ← Complement of H_0

.

Example: Stating the Null and Alternative Hypotheses

Write the claim as a mathematical sentence. State the null and alternative hypotheses and identify which represents the claim.

2. A car dealership announces that the mean time for an oil change is less than 15 minutes.

Solution:

H_0: $\mu \geq 15$ minutes ← Complement of H_a

H_a: $\mu < 15$ minutes ← Inequality condition (Claim)

.

Example: Stating the Null and Alternative Hypotheses

Write the claim as a mathematical sentence. State the null and alternative hypotheses and identify which represents the claim.

3. A company advertises that the mean life of its furnaces is more than 18 years

Solution:

H_0: $\mu \leq 18$ years ⟵ Complement of H_a

H_a: $\mu > 18$ years ⟵ Inequality condition (Claim)

.

Types of Errors

- No matter which hypothesis represents the claim, always begin the hypothesis test **assuming that the equality condition in the null hypothesis is true**.
- At the end of the test, one of two decisions will be made:
  - reject the null hypothesis
  - fail to reject the null hypothesis
- Because your decision is based on a sample, there is the possibility of making the wrong decision.

.

	Truth is Person Innocent	Truth is Person Guilty
Jury Decides Person Innocent	Correct Decision	Type II Error
Jury Decides Person Guilty	Type I Error	Correct Decision

Possible Outcomes and Types of Errors

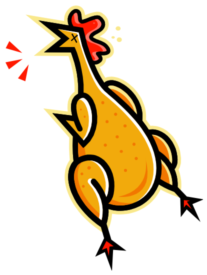
	Actual Truth of H_0	
Decision	H_0 is true	H_0 is false
Do not reject H_0	Correct Decision	**Type II Error**
Reject H_0	**Type I Error**	Correct Decision

- A **type I error** occurs if the null hypothesis is rejected when it is true.
- A **type II error** occurs if the null hypothesis is not rejected when it is false.

.

Example: Identifying Type I and Type II Errors

The USDA limit for salmonella contamination for chicken is 20%. A meat inspector reports that the chicken produced by a company exceeds the USDA limit. You perform a hypothesis test to determine whether the meat inspector's claim is true. When will a type I or type II error occur? Which is more serious? *(Source: United States Department of Agriculture)*



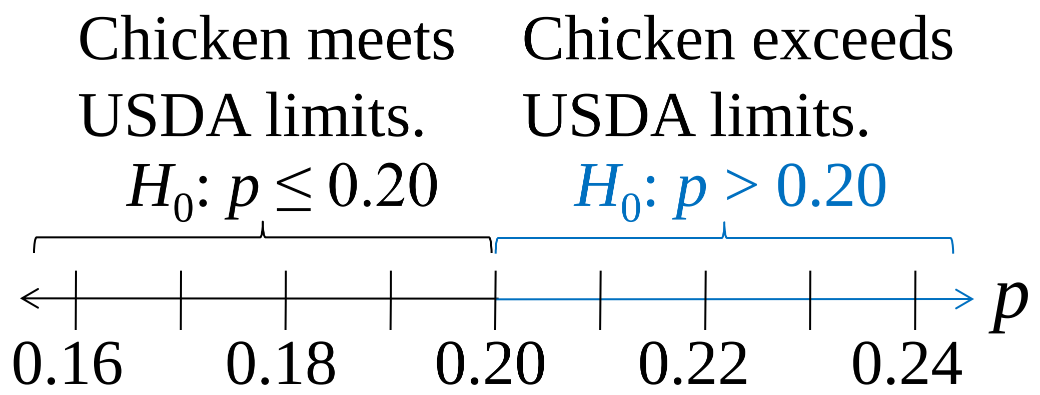
.

Solution: Identifying Type I and Type II Errors

Let p represent the proportion of chicken that is contaminated.

Hypotheses: H_0: $p \leq 0.2$

H_a: $p > 0.2$ (Claim)



.

Solution: Identifying Type I and Type II Errors

Hypotheses: H_0: $p \leq 0.2$

H_a: $p > 0.2$ (Claim)

A type I error is rejecting H_0 when it is true.

The actual proportion of contaminated chicken is less than or equal to 0.2, but you decide to reject H_0.

A type II error is failing to reject H_0 when it is false.

The actual proportion of contaminated chicken is greater than 0.2, but you do not reject H_0.

.

Solution: Identifying Type I and Type II Errors

Hypotheses: H_0: $p \leq 0.2$
H_a: $p > 0.2$ (Claim)

- With a type I error, you might create a health scare and hurt the sales of chicken producers who were actually meeting the USDA limits.
- With a type II error, you could be allowing chicken that exceeded the USDA contamination limit to be sold to consumers.
- A type II error could result in sickness or even death.

.

Level of Significance

Level of significance

- Your maximum allowable probability of making a type I error.
  - Denoted by α, the lowercase Greek letter alpha.
- By setting the level of significance at a small value, you are saying that you want the probability of rejecting a true null hypothesis to be small.
- Commonly used levels of significance:
  - $\alpha = 0.10$ $\alpha = 0.05$ $\alpha = 0.01$

.

Statistical Tests

- After stating the null and alternative hypotheses and specifying the level of significance, a random sample is taken from the population and sample statistics are calculated.
- The statistic that is compared with the parameter in the null hypothesis is called the **test statistic**.

Population parameter	Test statistic	Standardized test statistic
μ	$\bar{x}$	$z\ (n \geq 30)$ $t\ (n < 30)$
p	$\hat{p}$	z

.

***P*-values**

***P*-value** (or **probability value**)

- The probability, if the null hypothesis is true, of obtaining a sample statistic with a value as extreme or more extreme than the one determined from the sample data.
- Depends on the nature of the test.

.

Nature of the Test

- Three types of hypothesis tests
  - left-tailed test
  - right-tailed test
  - two-tailed test
- The type of test depends on the region of the sampling distribution that favors a rejection of H_0.
- This region is indicated by the alternative hypothesis.

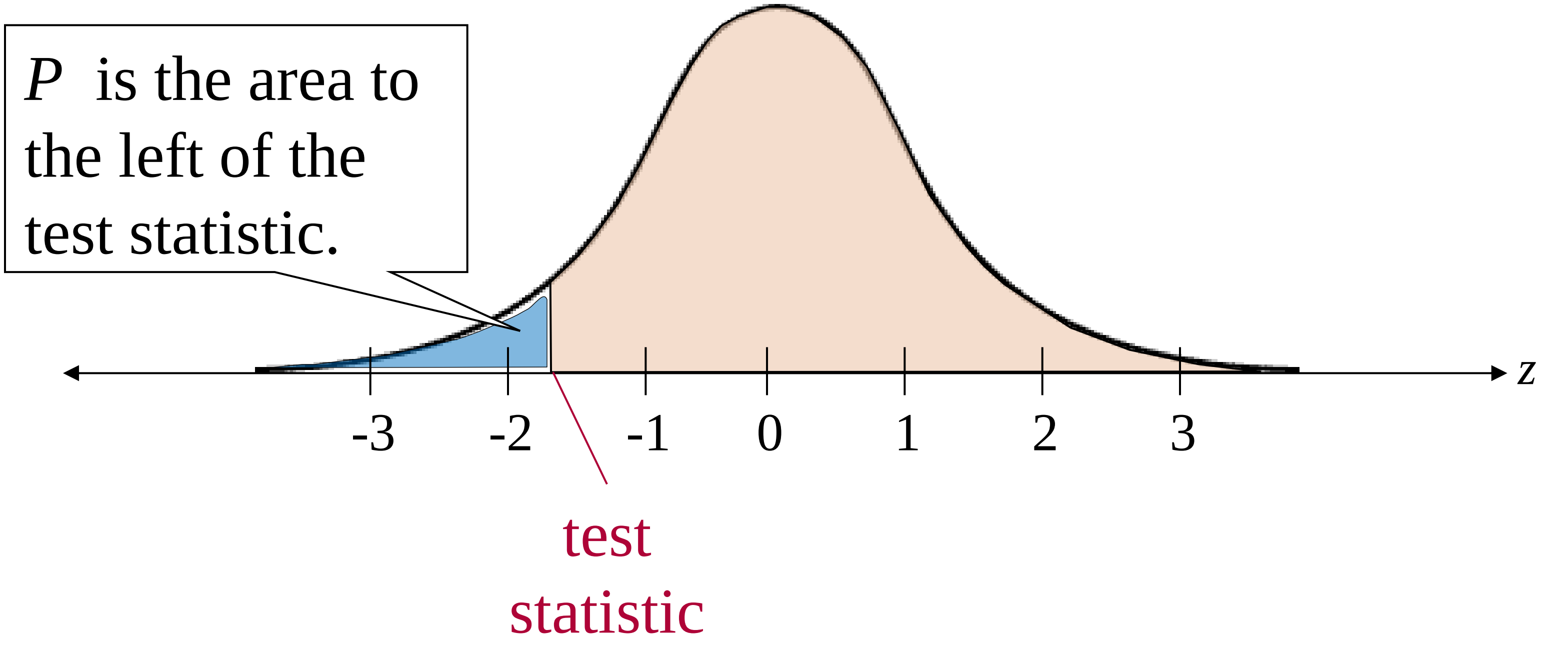
.

Left-tailed Test

- The alternative hypothesis H_1 contains the less-than inequality symbol (<).

$H_0: \mu \geq k$

$H_a: \mu < k$



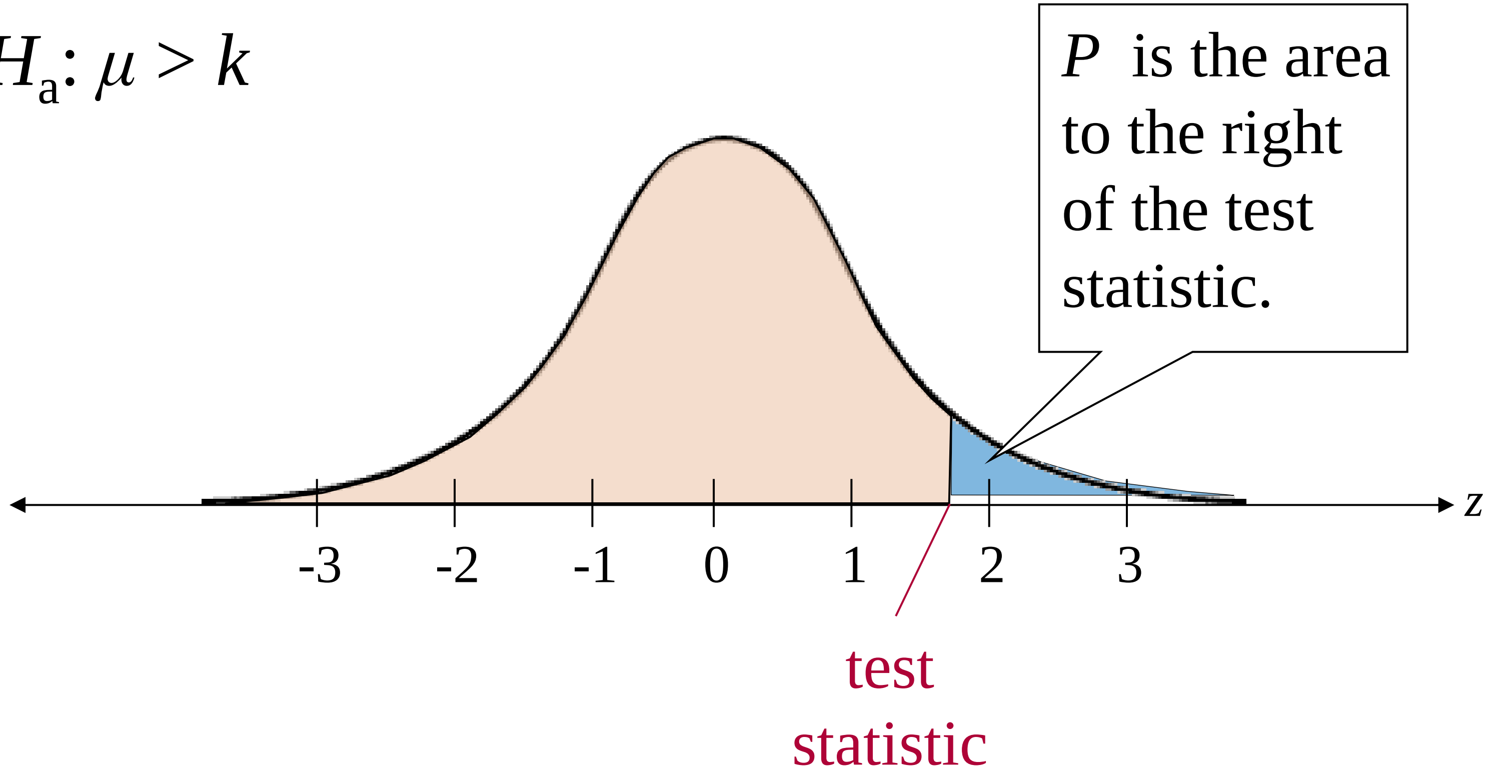
.

Right-tailed Test

- The alternative hypothesis H_1 contains the greater-than inequality symbol (>).

H_0: $\mu \leq k$

H_a: $\mu > k$



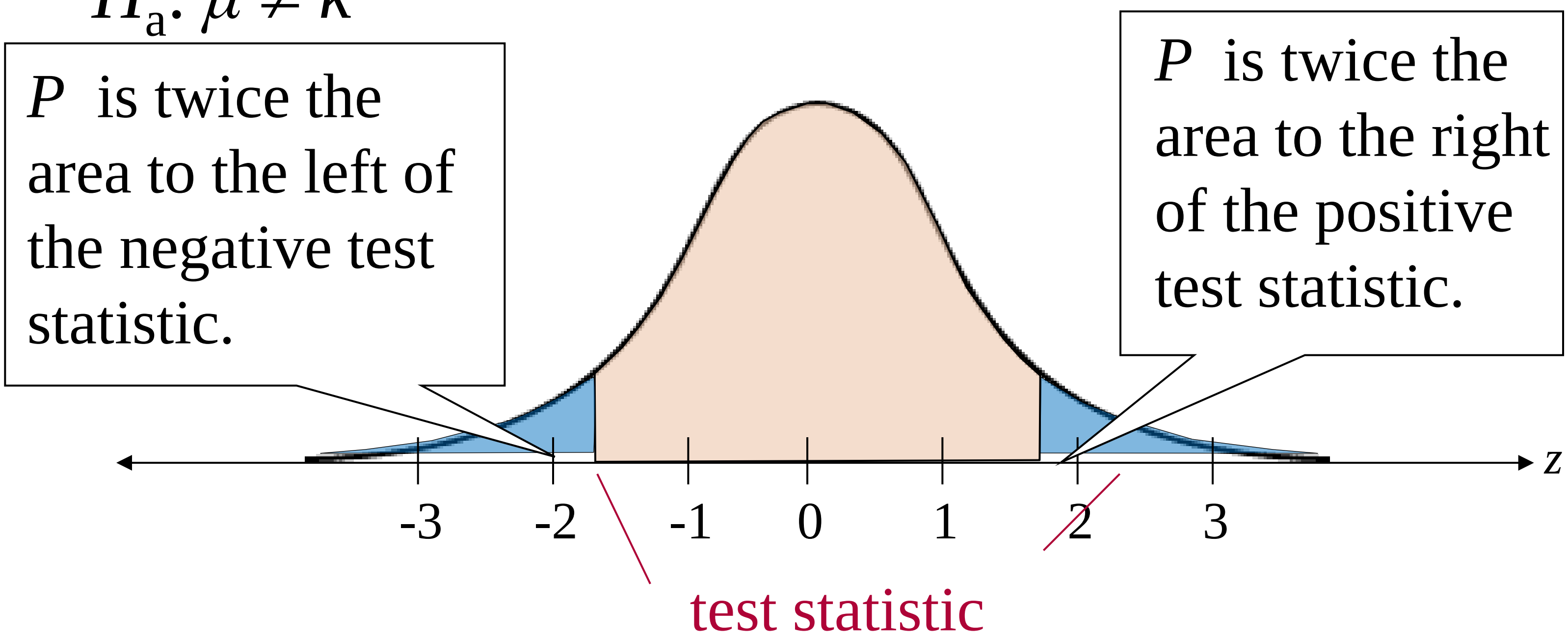
.

Two-tailed Test

- The alternative hypothesis H_1 contains the not equal inequality symbol ($\neq$). Each tail has an area of $\frac{1}{2}P$.

H_0: $\mu = k$

H_a: $\mu \neq k$



Example: Identifying The Nature of a Test

For each claim, state H_0 and H_a. Then determine whether the hypothesis test is a left-tailed, right-tailed, or two-tailed test. Sketch a normal sampling distribution and shade the area for the *P*-value.

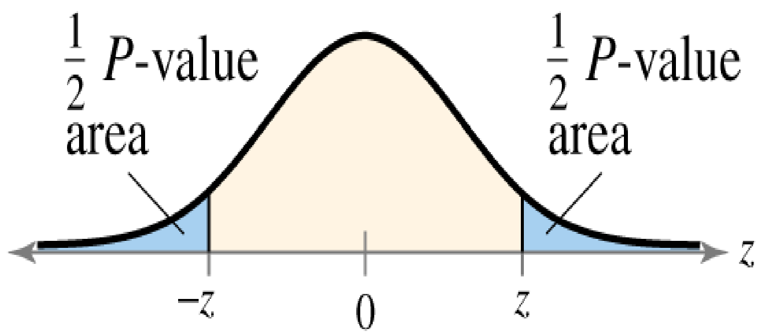
1. A school publicizes that the proportion of its students who are involved in at least one extracurricular activity is 61%.

Solution:

H_0: $p = 0.61$

H_a: $p \neq 0.61$

Two-tailed test



.

Example: Identifying The Nature of a Test

For each claim, state H_0 and H_1. Then determine whether the hypothesis test is a left-tailed, right-tailed, or two-tailed test. Sketch a normal sampling distribution and shade the area for the *P*-value.

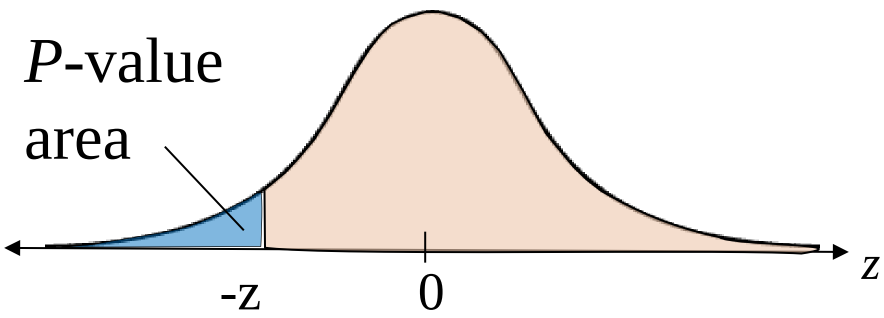
2. A car dealership announces that the mean time for an oil change is less than 15 minutes.

Solution:

$H_0: \mu \geq 15$ min

$H_a: \mu < 15$ min

Left-tailed test



.

Example: Identifying The Nature of a Test

For each claim, state H_0 and H_1. Then determine whether the hypothesis test is a left-tailed, right-tailed, or two-tailed test. Sketch a normal sampling distribution and shade the area for the P-value.

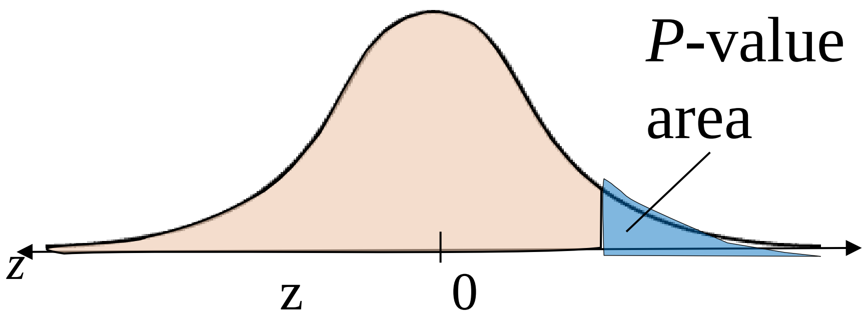
3. A company advertises that the mean life of its furnaces is more than 18 years.

Solution:

$H_0: \mu \le 18$ yr

$H_a: \mu > 18$ yr

Right-tailed test



.

Making a Decision

Decision Rule Based on *P*-value

- Compare the *P*-value with α.
  - If $P \le \alpha$, then reject H_0.
  - If $P > \alpha$, then fail to reject H_0.

	Claim	
Decision	Claim is H_0	Claim is H_a
Reject H_0	There is enough evidence to reject the claim	There is enough evidence to support the claim
Fail to reject H_0	There is not enough evidence to reject the claim	There is not enough evidence to support the claim

.

Steps for Hypothesis Testing

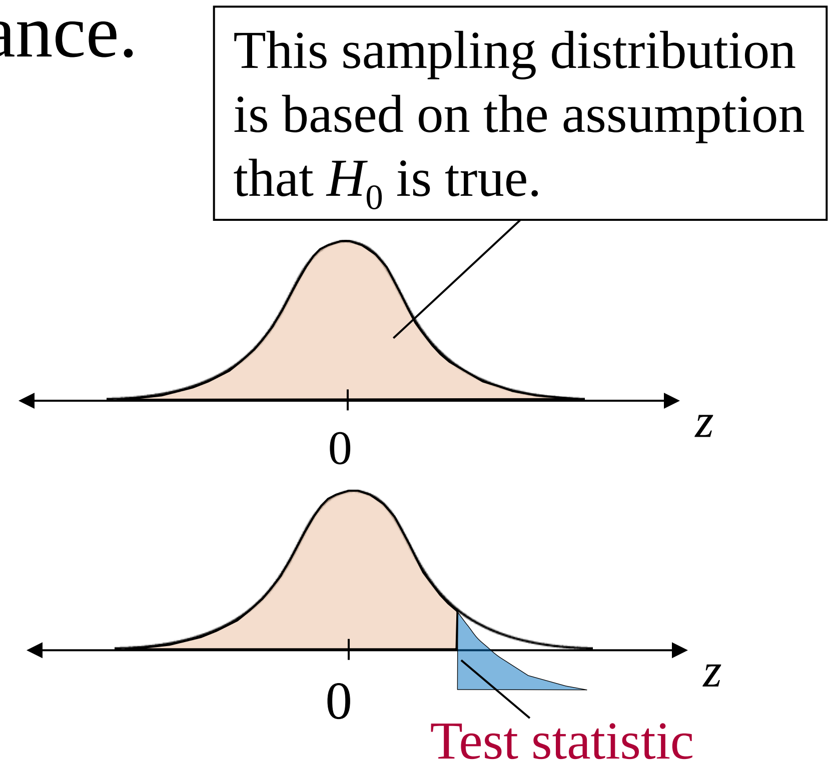
1. State the claim mathematically and verbally. Identify the null and alternative hypotheses.

   H_0: ? H_a: ?

2. Specify the level of significance.

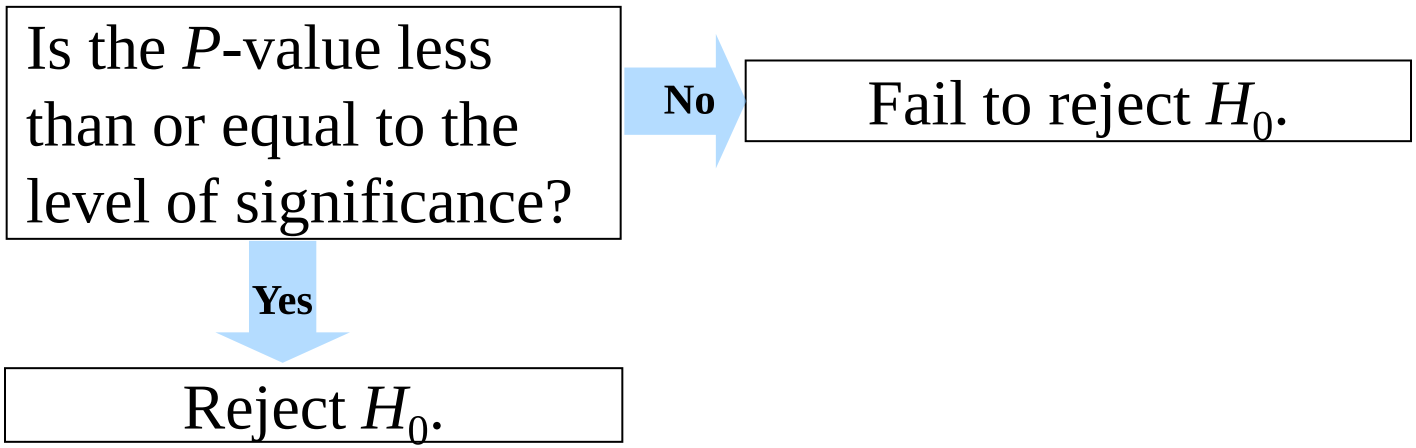
   $\alpha =$?

3. Determine the standardized sampling distribution and draw its graph.
4. Calculate the test statistic and its standardized value. Add it to your sketch.



Steps for Hypothesis Testing

5. Find the *P*-value.
6. Use the following decision rule.



7. Write a statement to interpret the decision in the context of the original claim.

Example: Hypothesis Testing Using *P*-values

In auto racing, a pit crew claims that its mean pit stop time (for 4 new tires and fuel) is less than 13 seconds. A random selection of 32 pit stop times has a sample mean of 12.9 seconds. Assume the population standard deviation is 0.19 second. Is there enough evidence to support the claim at $\alpha = 0.01$?

.

Solution: Hypothesis Testing Using *P*-values

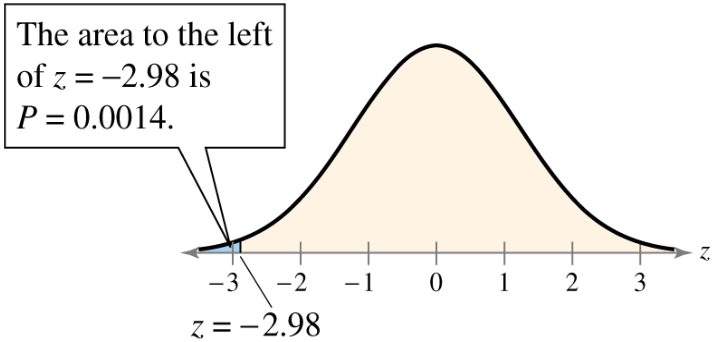
- H_0: $\mu \geq 13$ sec
- H_a: $\mu < 13$ sec
- $\alpha = 0.01$
- **Test Statistic:**

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

$$= \frac{12.9 - 13}{0.19/\sqrt{32}}$$

$$\approx -2.98$$

- ***P*-value**



Left-Tailed Test

- **Decision:** 0.0014 < 0.01

Reject H_0

At the 1% level of significance, you have sufficient evidence to conclude the mean pit stop time is less than 13 seconds.

.

Example: Hypothesis Testing Using *P*-values

According to a study, the mean cost of bariatric (weight loss) surgery is $21,500. You think this information is incorrect. You randomly select 25 bariatric surgery patients and find that the average cost for their surgeries is $20,695. The population standard deviation is known to be $2250 and the population is normally distributed. Is there enough evidence to support your claim at $\alpha = 0.05$?

Use a *P*-value.

.

Solution: Hypothesis Testing Using *P*-values

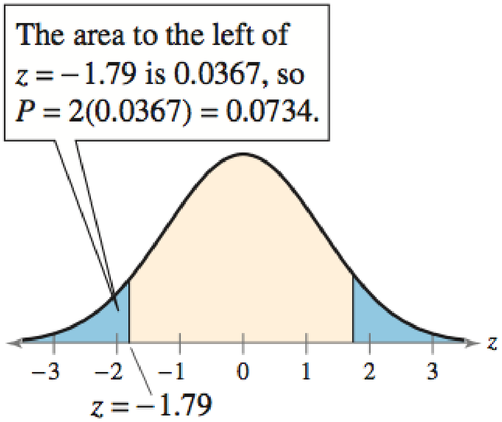
- H_0: μ = $21,500
- H_a: μ ≠ 21,500
- α = 0.05
- **Test Statistic:**

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

$$\approx \frac{20{,}695 - 21{,}500}{2250/\sqrt{25}}$$

$$\approx -1.79$$

- ***P*-value**



- **Decision:** 0.0734 > 0.05

Fail to reject H_0

At the 5% level of significance, there is not sufficient evidence to support the claim that the mean cost of bariatric surgery is different from $21,500.

Finding Critical Values in a *t*-Distribution

1. Identify the level of significance α.
2. Identify the degrees of freedom d.f. = $n - 1$.
3. Find the critical value(s) using Table 5 in Appendix B in the row with $n - 1$ degrees of freedom. If the hypothesis test is
   a. *left-tailed*, use "One Tail, α" column with a negative sign,
   b. *right-tailed*, use "One Tail, α" column with a positive sign,
   c. *two-tailed*, use "Two Tails, α" column with a negative and a positive sign.

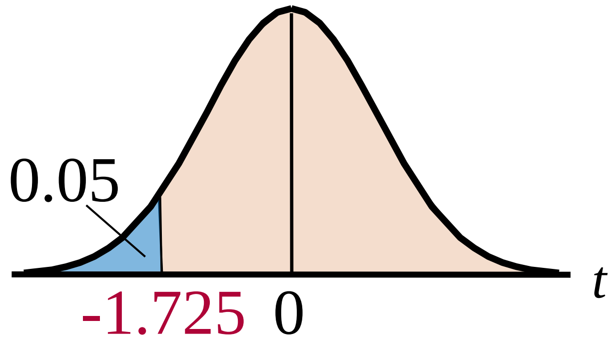
.

Example: Finding Critical Values for *t*

Find the critical value t_0 for a left-tailed test given $\alpha = 0.05$ and $n = 21$.

Solution:

- The degrees of freedom are d.f. = $n - 1 = 21 - 1 = 20$.
- Use Table 5.
- Look at α = 0.05 in the "One Tail, α" column.
- Because the test is left-tailed, the critical value is negative.



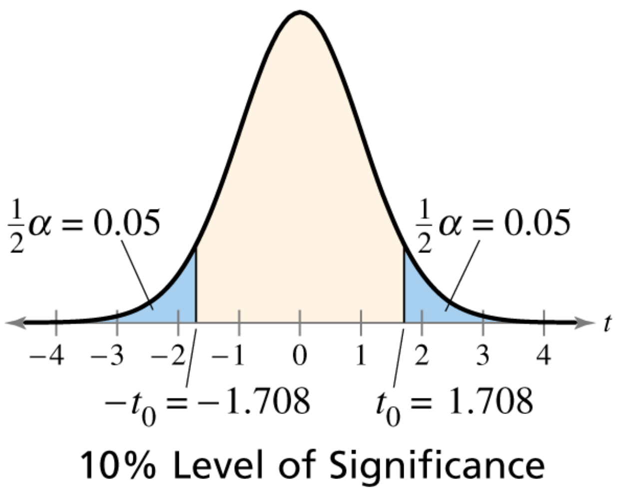
.

Example: Finding Critical Values for *t*

Find the critical values $-t_0$ and t_0 for a two-tailed test given $\alpha = 0.10$ and $n = 26$.

Solution:

- The degrees of freedom are d.f. = $n - 1 = 26 - 1 = 25$.
- Look at α = 0.10 in the "Two Tail, α" column.
- Because the test is two-tailed, one critical value is negative and one is positive.



.

t-Test for a Mean μ (*n* < 30, σ Unknown)

t-Test for a Mean

- A statistical test for a population mean.
- The *t*-test can be used when the population is normal or nearly normal, σ is unknown, and *n* < 30.
- The **test statistic** is the sample mean $\bar{x}$
- The **standardized test statistic** is *t*.

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

- The degrees of freedom are d.f. = $n - 1$.

.

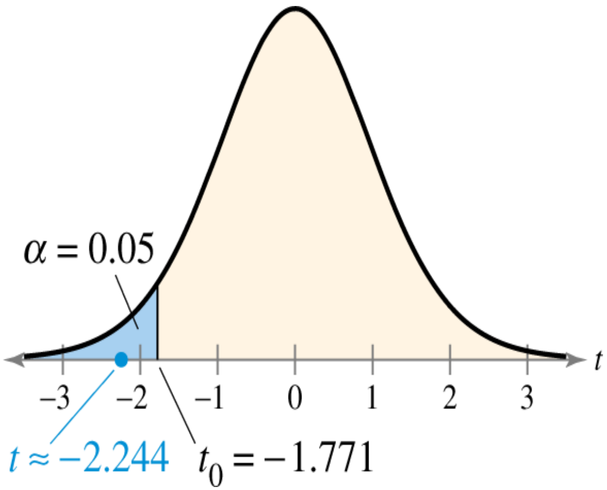
Example: Testing μ with a Small Sample

A used car dealer says that the mean price of a 2008 Honda CR-V is at least \$20,500. You suspect this claim is incorrect and find that a random sample of 14 similar vehicles has a mean price of \$19,850 and a standard deviation of \$1084. Is there enough evidence to reject the dealer's claim at $\alpha = 0.05$? Assume the population is normally distributed. *(Adapted from Kelley Blue Book)*

.

Solution: Testing μ with a Small Sample

- H_0: μ ≥ $20,500
- H_a: μ < $20,500
- α = 0.05
- df = 14 – 1 = 13
- Rejection Region:



- Test Statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{19{,}850 - 20{,}500}{1084/\sqrt{14}} \approx -2.244$$

- Decision:

Reject H_0

At the 0.05 level of significance, there is enough evidence to reject the claim that the mean price of a 2008 Honda CR-V is at least $20,500.

.

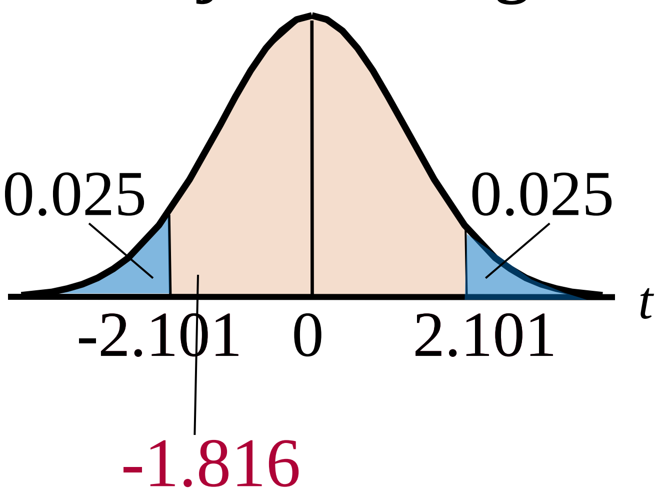
Example: Testing μ with a Small Sample

An industrial company claims that the mean pH level of the water in a nearby river is 6.8. You randomly select 19 water samples and measure the pH of each. The sample mean and standard deviation are 6.7 and 0.24, respectively. Is there enough evidence to reject the company's claim at $\alpha = 0.05$? Assume the population is normally distributed.

.

Solution: Testing μ with a Small Sample

- H_0: μ = 6.8
- H_a: μ ≠ 6.8
- α = 0.05
- df = 19 – 1 = 18
- Rejection Region:



- Test Statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{6.7 - 6.8}{0.24/\sqrt{19}} \approx -1.816$$

- Decision: Fail to reject H_0

At the 0.05 level of significance, there is not enough evidence to reject the claim that the mean pH is 6.8.

A department of motor vehicles office claims that the mean wait time is less than 14 minutes. A random sample of 10 people has a mean wait time of 13 minutes with a standard deviation of 3.5 minutes. At $\alpha = 0.10$, test the office's claim. Assume the population is normally distributed.

.

z-Test for a Population Proportion

z-Test for a Population Proportion

- A statistical test for a population proportion.
- Can be used when a binomial distribution is given such that $np \geq 5$ and $nq \geq 5$.
- The **test statistic** is the sample proportion $\hat{p}$.
- The **standardized test statistic** is z.

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

.

Example: Hypothesis Test for a Proportion

A research center claims that less than 50% of U.S. adults have accessed the Internet over a wireless network with a laptop computer. In a random sample of 100 adults, 39% say they have accessed the Internet over a wireless network with a laptop computer. At $\alpha = 0.01$, is there enough evidence to support the researcher's claim? *(Adopted from Pew Research Center)*

Solution:

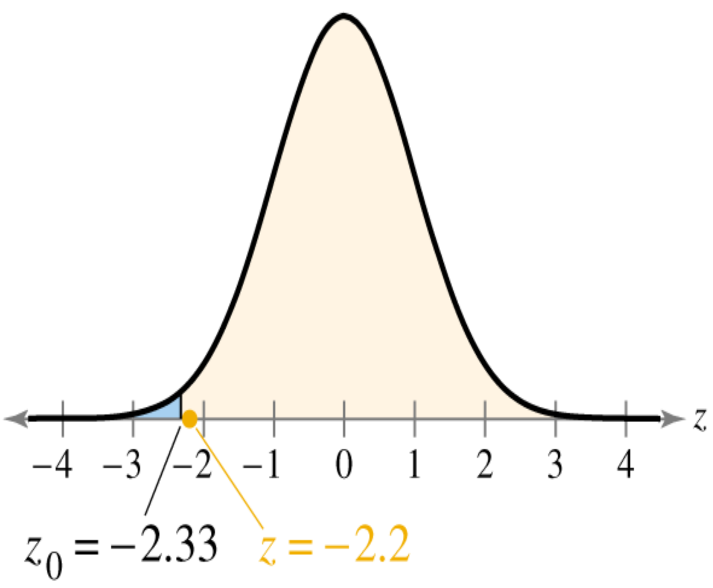
- Verify that $np \geq 5$ and $nq \geq 5$.

$$np = 100(0.50) = 50 \quad \text{and} \quad nq = 100(0.50) = 50$$

.

Solution: Hypothesis Test for a Proportion

- H_0: $p \geq 0.5$
- H_a: $p < 0.5$
- $\alpha =$ 0.01
- **Rejection Region:**



- **Test Statistic**

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.39 - 0.5}{\sqrt{(0.5)(0.5)/100}}$$

$$= -2.2$$

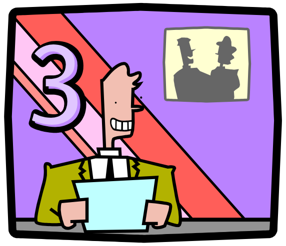
- **Decision:** **Fail to reject H_0**

At the 1% level of significance, there is not enough evidence to support the claim that less than 50% of U.S. adults have accessed the Internet over a wireless network with a laptop computer.

.

Example: Hypothesis Test for a Proportion

The Research Center claims that 25% of college graduates think a college degree is not worth the cost. You decide to test this claim and ask a random sample of 200 college graduates whether they think a college is not worth the cost. Of those surveyed, 21% reply yes. At $\alpha = 0.10$ is there enough evidence to reject the claim?



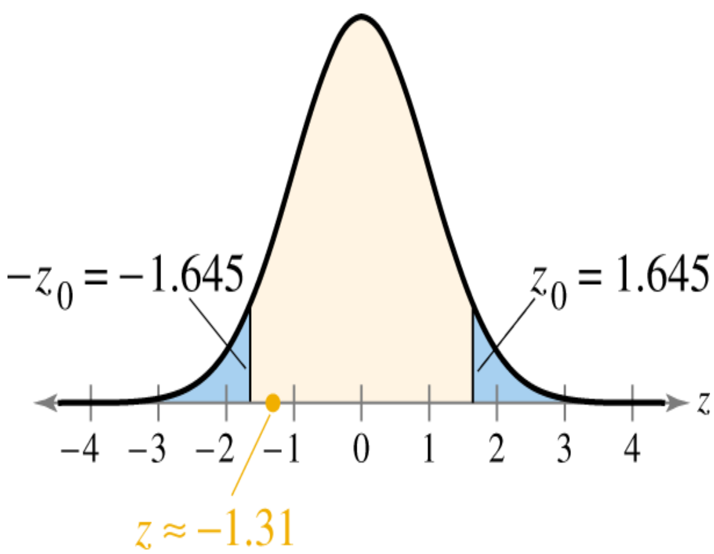
Solution:

- Verify that $np \geq 5$ and $nq \geq 5$.

$np = 200(0.25) = 50$ and $nq = 200\,(0.75) = 150$

Solution: Hypothesis Test for Proportions

- H_0: $p = 0.25$
- H_a: $p \neq 0.25$
- $\alpha = 0.10$
- **Rejection Region:**



- **Test Statistic**

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.21 - 0.25}{\sqrt{(0.25)(0.75)/200}}$$

$$= -1.31$$

- **Decision:** **Fail to reject H_0**

At the 10% level of significance, there is not enough evidence to reject the claim that 25% of college graduates think a college degree is not worth the cost.