ISE 315: Engineering Statistics

Lecture 8: Introduction to Hypothesis Testing

Instructor: Mansur M. Arief, PhD
Industrial and Systems Engineering, KFUPM

Office: 22-219 — Email: mansur.arief@kfupm.edu.sa

Based on Montgomery & Runger, Applied Statistics and Probability for Engineers, 6th Ed.

Lecture 8

Introduction to Hypothesis Testing (Chapter 9)

Lecture 8 Outline

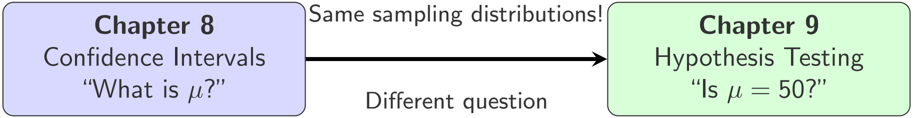
- Introduction to Hypothesis Testing (Chapter 9)
  - From confidence intervals to hypothesis tests
  - Statistical hypotheses: H_0 vs H_1
  - Type I and Type II errors
  - Test statistics and critical regions
- Quiz #2 (30 minutes)

Chapter 9

Tests of Hypotheses for a Single Sample

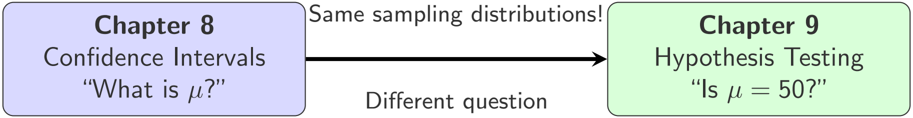
From Confidence Intervals to Hypothesis Testing

The Big Picture



From Confidence Intervals to Hypothesis Testing

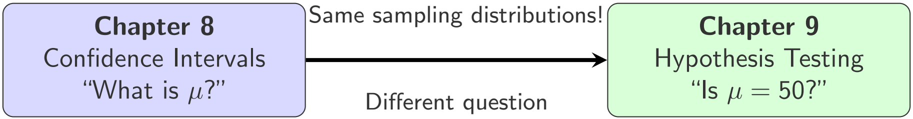
The Big Picture



- **Chapter 8 (CI):** Estimate an unknown parameter
  - "The true mean viscosity is between 116.08 and 123.92 cP"

From Confidence Intervals to Hypothesis Testing

The Big Picture



- **Chapter 8 (CI):** Estimate an unknown parameter
  - “The true mean viscosity is between 116.08 and 123.92 cP”
- **Chapter 9 (HT):** Test a claim about a parameter
  - “Is the true mean viscosity equal to 120 cP?”

Motivating Example: Crude Oil Viscosity

Scenario: A petroleum engineer measures crude oil viscosity.

- Sample: $n = 64$ measurements, $\bar{x} = 120$ cP, known $\sigma = 16$ cP

Motivating Example: Crude Oil Viscosity

Scenario: A petroleum engineer measures crude oil viscosity.

- Sample: $n = 64$ measurements, $\bar{x} = 120$ cP, known $\sigma = 16$ cP

Chapter 8 Question: What is the true mean viscosity?

- 95% CI: $(116.08, 123.92)$ cP

Motivating Example: Crude Oil Viscosity

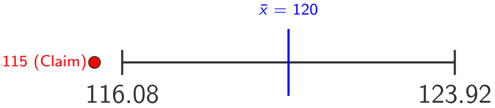
Scenario: A petroleum engineer measures crude oil viscosity.

- Sample: $n = 64$ measurements, $\bar{x} = 120$ cP, known $\sigma = 16$ cP

Chapter 8 Question: What is the true mean viscosity?

- 95% CI: $(116.08, 123.92)$ cP

Chapter 9 Question: The supplier claims the mean viscosity is 115 cP. Is this claim supported by our data?



Motivating Example: Crude Oil Viscosity

Scenario: A petroleum engineer measures crude oil viscosity.

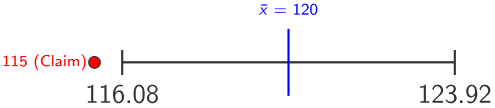
- Sample: $n = 64$ measurements, $\bar{x} = 120$ cP, known $\sigma = 16$ cP

Chapter 8 Question: What is the true mean viscosity?

- 95% CI: $(116.08, 123.92)$ cP

Chapter 9 Question: The supplier claims the mean viscosity is 115 cP. Is this claim supported by our data?

- 115 is **outside** our 95% CI! (the claim might be false?)



What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

The Process:

What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

The Process:

1. State a **hypothesis** (claim) about the parameter

What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

The Process:

1. State a **hypothesis** (claim) about the parameter
2. Collect sample data

What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

The Process:

1. State a **hypothesis** (claim) about the parameter
2. Collect sample data
3. Calculate a **test statistic**

What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

The Process:

1. State a **hypothesis** (claim) about the parameter
2. Collect sample data
3. Calculate a **test statistic**
4. Make a **decision**: Reject or fail to reject the hypothesis

What is Hypothesis Testing?

Definition: A procedure for making decisions about population parameters based on sample data.

The Process:

1. State a **hypothesis** (claim) about the parameter
2. Collect sample data
3. Calculate a **test statistic**
4. Make a **decision**: Reject or fail to reject the hypothesis

Key Idea: We use probability to quantify the evidence against the hypothesis.

The Two Hypotheses

H_0 and H_1

Null Hypothesis H_0:

- The “default” or “status quo”
- What we assume is true
- Usually contains “$=$”
- Example: $H_0 : \mu = 115$

The Two Hypotheses

H_0 and H_1

Null Hypothesis H_0:

- The "default" or "status quo"
- What we assume is true
- Usually contains "$=$"
- Example: $H_0 : \mu = 115$

Alternative Hypothesis H_1:

- The "research hypothesis"
- What we want to prove
- Contains "$\neq$", "$<$", or "$>$"
- Example: $H_1 : \mu \neq 115$

The Two Hypotheses

H_0 and H_1

Null Hypothesis H_0:

- The "default" or "status quo"
- What we assume is true
- Usually contains "$=$"
- Example: $H_0 : \mu = 115$

Alternative Hypothesis H_1:

- The "research hypothesis"
- What we want to prove
- Contains "$\neq$", "$<$", or "$>$"
- Example: $H_1 : \mu \neq 115$

Important: We never "accept" H_0.
We either **reject** H_0 or **fail to reject** H_0.

Three Types of Hypothesis Tests

Type	**Hypotheses**	**Example**
Two-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$	Is μ different from 115?

Three Types of Hypothesis Tests

Type	**Hypotheses**	**Example**
Two-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$	Is μ different from 115?
Left-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$	Is μ less than 115?

Three Types of Hypothesis Tests

Type	Hypotheses	Example
Two-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$	Is μ different from 115?
Left-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$	Is μ less than 115?
Right-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$	Is μ greater than 115?

Three Types of Hypothesis Tests

Type	**Hypotheses**	**Example**
Two-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$	Is μ different from 115?
Left-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$	Is μ less than 115?
Right-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$	Is μ greater than 115?

Can you guess the connection to one- or two-sided CIs?

Type I and Type II Errors

Two Ways to Be Wrong

	Truth (Unknown)	
Decision	H_0 is True	H_0 is False

Type I and Type II Errors

Two Ways to Be Wrong

	Truth (Unknown)	
Decision	H_0 is True	H_0 is False
Reject H_0	**Type I Error** (α)	Correct!

Type I and Type II Errors

Two Ways to Be Wrong

Decision	Truth (Unknown)	
	H_0 is True	H_0 is False
Reject H_0	**Type I Error** (α)	Correct!
Fail to Reject H_0	Correct!	**Type II Error** (β)

Type I and Type II Errors

Two Ways to Be Wrong

Decision	Truth (Unknown)	
	H_0 is True	H_0 is False
Reject H_0	**Type I Error** (α)	Correct!
Fail to Reject H_0	Correct!	**Type II Error** (β)

- **Type I Error (α):** Rejecting H_0 when it's actually true
  - "False positive" or "false alarm", probability $= \alpha$

Type I and Type II Errors

Two Ways to Be Wrong

Decision	**Truth (Unknown)**	
	H_0 is True	H_0 is False
Reject H_0	**Type I Error** (α)	Correct!
Fail to Reject H_0	Correct!	**Type II Error** (β)

- **Type I Error (α):** Rejecting H_0 when it's actually true
  - "False positive" or "false alarm", probability $= \alpha$
- **Type II Error (β):** Failing to reject H_0 when it's actually false
  - "False negative" or "missed detection", probability $= \beta$

Engineering Analogy: Quality Control

Type I vs Type II Errors

Scenario: Testing if a batch of parts meets specifications.

- H_0: Batch is **good** (meets specs)
- H_1: Batch is **bad** (doesn't meet specs)

Engineering Analogy: Quality Control

Type I vs Type II Errors

Scenario: Testing if a batch of parts meets specifications.

- H_0: Batch is **good** (meets specs)
- H_1: Batch is **bad** (doesn't meet specs)

Type I Error
(Reject good batch)

- Waste parts, rework
- Cost: $$

Engineering Analogy: Quality Control

Type I vs Type II Errors

Scenario: Testing if a batch of parts meets specifications.

- H_0: Batch is **good** (meets specs)
- H_1: Batch is **bad** (doesn't meet specs)

Type I Error
(Reject good batch)

- Waste parts, rework
- Cost: \$\$

Type II Error
(Accept bad batch)

- Recalls, reputation damage
- Cost: \$\$\$\$

Engineering Analogy: Quality Control

Type I vs Type II Errors

Scenario: Testing if a batch of parts meets specifications.

- H_0: Batch is **good** (meets specs)
- H_1: Batch is **bad** (doesn't meet specs)

Type I Error
(Reject good batch)

- Waste parts, rework
- Cost: $$

Type II Error
(Accept bad batch)

- Recalls, reputation damage
- Cost: $$$$

Trade-off: Reducing α typically increases β (and vice versa).

Significance Level α

Controlling Type I Error

Definition: The significance level α is the maximum probability of Type I error we're willing to accept.

Significance Level α

Controlling Type I Error

Definition: The significance level α is the maximum probability of Type I error we're willing to accept.

Common choices:

- $\alpha = 0.10$ (10%) – Less stringent
- $\alpha = 0.05$ (5%) – Most common
- $\alpha = 0.01$ (1%) – More stringent

Significance Level α

Controlling Type I Error

Definition: The significance level α is the maximum probability of Type I error we're willing to accept.

Common choices:

- $\alpha = 0.10$ (10%) – Less stringent
- $\alpha = 0.05$ (5%) – Most common
- $\alpha = 0.01$ (1%) – More stringent

You have seen this in Chapter 8: $\alpha = 0.05$ for hypothesis test $\leftrightarrow$ 95% CI

The Test Statistic

Measuring Evidence Against H_0

Idea: Convert sample data into a single number that measures how far the data is from what H_0 predicts.

The Test Statistic

Measuring Evidence Against H_0

Idea: Convert sample data into a single number that measures how far the data is from what H_0 predicts.

For testing μ when σ is known: use the Z-test statistic: $Z_0 = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$

- $\bar{X}$ = sample mean (what we observed)
- μ_0 = hypothesized value (from H_0)
- $\sigma/\sqrt{n}$ = standard error

The Test Statistic

Measuring Evidence Against H_0

Idea: Convert sample data into a single number that measures how far the data is from what H_0 predicts.

For testing μ when σ is known: use the Z-test statistic: $Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

- $\bar{X}$ = sample mean (what we observed)
- μ_0 = hypothesized value (from H_0)
- $\sigma/\sqrt{n}$ = standard error

The Test Statistic

Measuring Evidence Against H_0

Idea: Convert sample data into a single number that measures how far the data is from what H_0 predicts.

For testing μ when σ is known: use the Z-test statistic: $Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

- $\bar{X}$ = sample mean (what we observed)
- μ_0 = hypothesized value (from H_0)
- $\sigma / \sqrt{n}$ = standard error

Interpretation: Z_0 tells us how many standard errors $\bar{X}$ is from μ_0.

Example: Testing the Supplier's Claim

Crude Oil Viscosity

Given: $n = 64$, $\bar{x} = 120$ cP, $\sigma = 16$ cP
Supplier claims: $\mu = 115$ cP

Example: Testing the Supplier's Claim

Crude Oil Viscosity

Given: $n = 64$, $\bar{x} = 120$ cP, $\sigma = 16$ cP
Supplier claims: $\mu = 115$ cP

Hypotheses: $H_0 : \mu = 115$ vs $H_1 : \mu \neq 115$

Example: Testing the Supplier's Claim

Crude Oil Viscosity

Given: $n = 64$, $\bar{x} = 120$ cP, $\sigma = 16$ cP
Supplier claims: $\mu = 115$ cP

Hypotheses: $H_0 : \mu = 115$ vs $H_1 : \mu \neq 115$
Test statistic: $Z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{120-115}{16/\sqrt{64}} = \frac{5}{2} = 2.5$

Example: Testing the Supplier's Claim

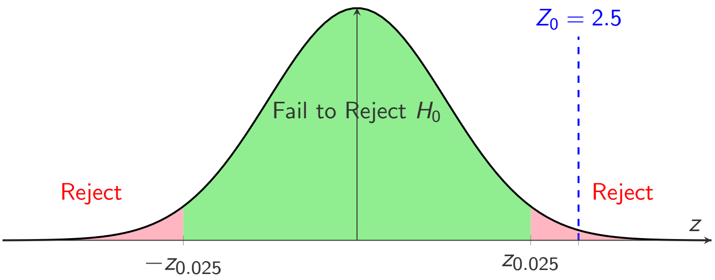
Crude Oil Viscosity

Given: $n = 64$, $\bar{x} = 120$ cP, $\sigma = 16$ cP
Supplier claims: $\mu = 115$ cP

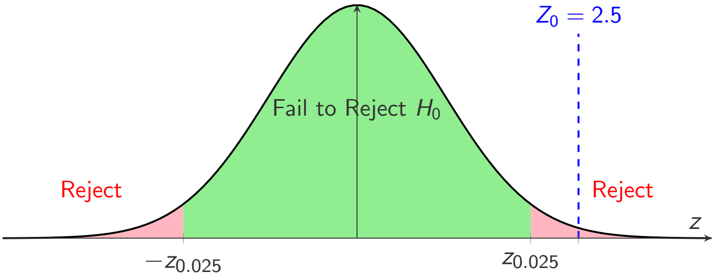
Hypotheses: $H_0 : \mu = 115$ vs $H_1 : \mu \neq 115$
Test statistic: $Z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{120-115}{16/\sqrt{64}} = \frac{5}{2} = 2.5$
Question: Is $Z_0 = 2.5$ large enough to reject H_0?

We need a **decision rule**... (next slide!)

Critical Region and Decision Rule



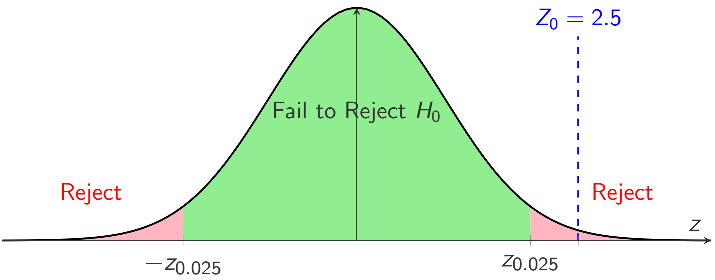
Critical Region and Decision Rule



Decision Rule (for $\alpha = 0.05$, two-tailed): Reject H_0 if $|Z_0| > z_{0.025} = 1.96$

- Since $|2.5| > 1.96$, we **reject** H_0

Critical Region and Decision Rule



Decision Rule (for $\alpha = 0.05$, two-tailed): Reject H_0 if $|Z_0| > z_{0.025} = 1.96$

- Since $|2.5| > 1.96$, we **reject** H_0

Conclusion: The data provides sufficient evidence that the claim is **false**.

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against
- H_1 **(alternative):** What we're trying to prove

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against
- H_1 **(alternative):** What we're trying to prove
- **Type I error (α):** Rejecting a true H_0 (false positive)

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against
- H_1 **(alternative):** What we're trying to prove
- **Type I error (α):** Rejecting a true H_0 (false positive)
- **Type II error (β):** Failing to reject a false H_0 (false negative)

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against
- H_1 **(alternative):** What we're trying to prove
- **Type I error (α):** Rejecting a true H_0 (false positive)
- **Type II error (β):** Failing to reject a false H_0 (false negative)
- **Test statistic:** Measures evidence against H_0

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against
- H_1 **(alternative):** What we're trying to prove
- **Type I error (α):** Rejecting a true H_0 (false positive)
- **Type II error (β):** Failing to reject a false H_0 (false negative)
- **Test statistic:** Measures evidence against H_0
- **Decision:** Compare test statistic to critical value

Chapter 9: Key Takeaways So Far

- **Hypothesis testing** is about making decisions based on sample data
- H_0 **(null):** The default claim we test against
- H_1 **(alternative):** What we're trying to prove
- **Type I error (**α**):** Rejecting a true H_0 (false positive)
- **Type II error (**β**):** Failing to reject a false H_0 (false negative)
- **Test statistic:** Measures evidence against H_0
- **Decision:** Compare test statistic to critical value

Next lecture: More examples, P-values, and tests for σ unknown

Coming Up in Chapter 9...

Parameter	Test Statistic
Mean μ (σ known)	$Z_0 = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$
Mean μ (σ unknown)	$T_0 = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$
Variance σ^2	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$
Proportion p	$Z_0 = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$

Coming Up in Chapter 9...

Parameter	Test Statistic
Mean μ (σ known)	$Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$
Mean μ (σ unknown)	$T_0 = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$
Variance σ^2	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$
Proportion p	$Z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

Notice: Same distributions as confidence intervals!