

Statistics for Economists II

Lecture 9: Hypothesis Testing

Güneş A. Aşık

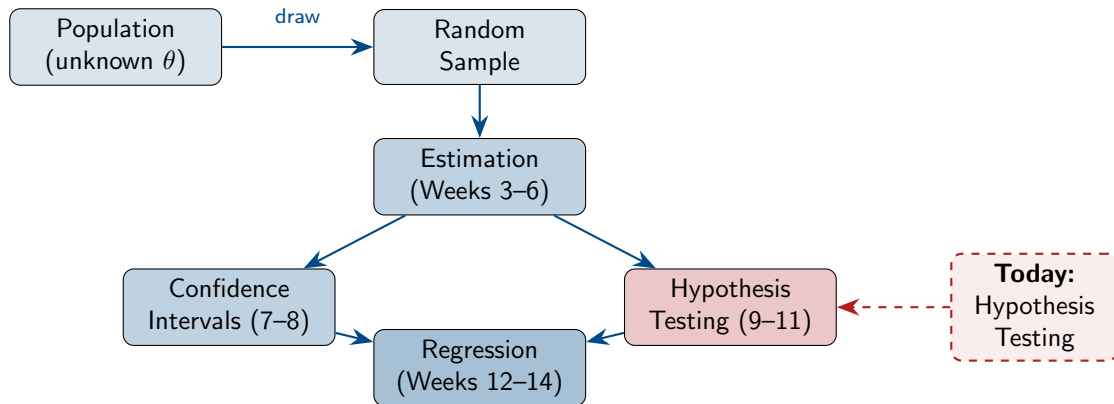
Department of Economics
Middle East Technical University

Spring 2025–2026

Roadmap for Today

- 1 From Estimation to Hypothesis Testing
- 2 The Logic of Hypothesis Testing
- 3 Types of Errors
- 4 One-Sided and Two-Sided Tests
- 5 The P -Value Approach
- 6 CI and Hypothesis Testing: Two Sides of the Same Coin
- 7 Z-Test for the Mean (Known σ)
- 8 Example 5.1: Phone Bills (One-Sided Test)
- 9 Example 5.2: TV Sets (Two-Sided Test)
- 10 Economics Applications
- 11 A Summary of the Three Approaches
- 12 Practice Problems
- 13 Step-by-Step Procedure
- 14 Common Mistakes and Warnings
- 15 Key Takeaways

Where We Are in the Course



Two Fundamental Questions in Inference

Estimation (Weeks 3–8)

“What is the value of θ ?”

- Point estimates: $\hat{\theta}$
- Interval estimates: $(\hat{\theta}_L, \hat{\theta}_U)$
- We *quantify* the parameter

Hypothesis Testing (Weeks 9–11)

“Is this value of θ plausible?”

- Someone claims $\theta = \theta_0$
- We ask: does the data support this?
- We make a **yes/no decision**

Estimation is like asking “What is the income per capita of this country?”

Hypothesis testing is like asking “Could the income per capita be higher than USD 15,000?”

Why Do Economists Need Hypothesis Testing?

- 1 **Policy evaluation:** Did the minimum wage increase change average employment?

Why Do Economists Need Hypothesis Testing?

- ① **Policy evaluation:** Did the minimum wage increase change average employment?
- ② **Market efficiency:** Is the average stock return different from zero?

Why Do Economists Need Hypothesis Testing?

- ① **Policy evaluation:** Did the minimum wage increase change average employment?
- ② **Market efficiency:** Is the average stock return different from zero?
- ③ **Consumer behavior:** Has average household spending changed after a tax reform?

Why Do Economists Need Hypothesis Testing?

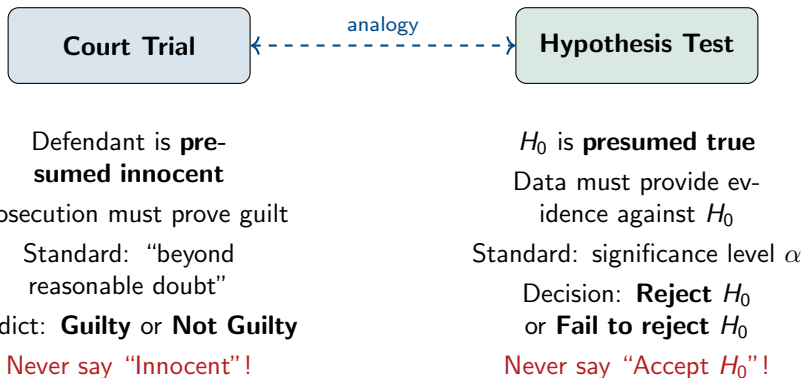
- ① **Policy evaluation:** Did the minimum wage increase change average employment?
- ② **Market efficiency:** Is the average stock return different from zero?
- ③ **Consumer behavior:** Has average household spending changed after a tax reform?
- ④ **Central banking:** Is inflation significantly above the 2% target?

Why Do Economists Need Hypothesis Testing?

- ① **Policy evaluation:** Did the minimum wage increase change average employment?
- ② **Market efficiency:** Is the average stock return different from zero?
- ③ **Consumer behavior:** Has average household spending changed after a tax reform?
- ④ **Central banking:** Is inflation significantly above the 2% target?

In each case, we have a **claim** about a population parameter. We collect sample data and ask: *Is the evidence strong enough to reject this claim?* This is the essence of hypothesis testing.

The Court Trial Analogy



Just as a "Not Guilty" verdict does not mean the defendant is innocent—it means the evidence was insufficient—"Fail to reject H_0 " does not mean H_0 is true. It means the data did not provide enough evidence against it.

The Four Components of a Hypothesis Test

- ➊ **Null hypothesis (H_0):** The hypothesis to be tested—often representing the *status quo* or *no effect*
- ➋ **Alternative hypothesis (H_a or H_1):** The research hypothesis we seek evidence to support
- ➌ **Test statistic:** A function of sample measurements on which the decision is based
- ➍ **Rejection region (RR):** The set of test statistic values for which we reject H_0

Decision Rule

- If test statistic $\in$ RR $\Rightarrow$ **Reject H_0**
- If test statistic $\notin$ RR $\Rightarrow$ **Fail to reject H_0**

The TWO DEADLY SINS of Hypothesis Testing

DEADLY SIN #1: Never Say “Accept H_0 ”

When H_a is not sufficiently supported by the data, use one of these phrases:

- *Fail to reject H_0*
- *Insufficient evidence to reject H_0*

DEADLY SIN #2: Never Include H_a in Your Concluding Statement

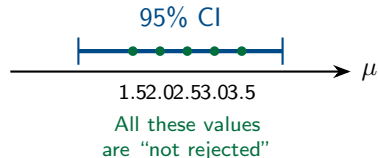
Avoid creative wording when making technical statistical statements. Your conclusion is always about H_0 —whether you reject it or fail to reject it.

Why “Fail to Reject” and Not “Accept”?

Consider testing $H_0 : \mu = 2$ vs. $H_1 : \mu \neq 2$.

Suppose a 95% CI is (1.5, 3.2).

- $\mu = 2$ is in the CI $\Rightarrow$ cannot reject $H_0 : \mu = 2$
- But $\mu = 1.8$ is also in the CI!
- And so is $\mu = 2.5$, $\mu = 3.0$, ...



Many values of μ are consistent with the data.

We cannot single out $\mu = 2$ and declare it “accepted.”

“Failing to reject H_0 ” is like a jury verdict of “Not Guilty.” It does not prove innocence—it only says the evidence was not strong enough to convict.

Type I and Type II Errors

	Fail to Reject H_0	Reject H_0
H_0 is True	Correct decision	Type I Error (α)
H_0 is False	Type II Error (β)	Correct decision (Power = $1 - \beta$)

Type I Error (α)

Rejecting H_0 when it is **true**.

Court analogy: Convicting an *innocent* person.

Probability = α (significance level)

Type II Error (β)

Failing to reject H_0 when it is **false**.

Court analogy: Letting a *guilty* person go free.

Probability = β

The Trade-Off Between α and β

- α , the significance level is set by the researcher in advance.
- For a fixed sample size n , reducing α **increases** β (and vice versa)
- The only way to reduce **both** is to increase n
- In practice, we fix α typically at 0.01, 0.05, or 0.10 and try to minimize β

Why do we control α ?

Because Type I error is usually considered more “serious”—like wrongful conviction.

Think of a smoke detector:

- **Type I:** Alarm when there is no fire (false alarm, annoying)
- **Type II:** No alarm when there *is* a fire (dangerous!)

Which error you tolerate more depends on the **context**. In medical testing, Type II errors (missing a disease) can be deadly.

Significance Level α : Setting the Standard of Proof

- The **significance level** α defines the unlikely values of the test statistic if H_0 is true
- It defines the **rejection region** of the sampling distribution
- α is selected by the researcher **before** conducting the test

α	Confidence Level ($1 - \alpha$)	Interpretation
0.10	90%	Less strict (easier to reject H_0)
0.05	95%	Standard in most research
0.01	99%	Very strict (hard to reject H_0)

Recall that the typical confidence levels—99%, 95%, and 90%—correspond to $\alpha = 0.01, 0.05, 0.10$. Confidence intervals and hypothesis tests use the **same** α

One-Sided vs. Two-Sided Tests

Upper-Tail (Right)

$$H_0 : \mu \leq \mu_0$$

$$H_1 : \mu > \mu_0$$

Reject if $Z > z_\alpha$

Lower-Tail (Left)

$$H_0 : \mu \geq \mu_0$$

$$H_1 : \mu < \mu_0$$

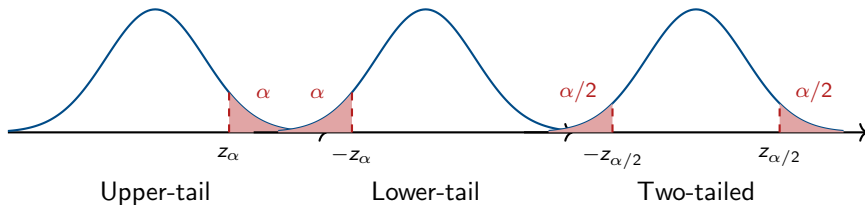
Reject if $Z < -z_\alpha$

Two-Tailed

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Reject if $|Z| > z_{\alpha/2}$



The P -Value: An Alternative to the Rejection Region

Definition

The **p -value** is the probability of obtaining a test statistic as extreme as, or more extreme than, the observed sample value, *given that H_0 is true*.

- It is also called the **observed level of significance**
- It represents the **smallest value of α** for which H_0 can be rejected based on the observed data

Decision Rule Using P -Value

- If $p\text{-value} < \alpha \Rightarrow$ **Reject H_0**
- If $p\text{-value} \geq \alpha \Rightarrow$ **Fail to reject H_0**

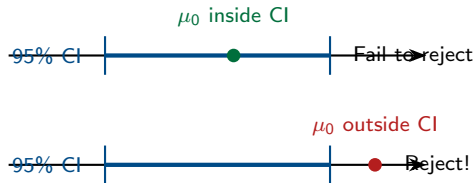
The p -value answers: “If H_0 were true, how surprised should we be by this data?” A small p -value means the data is **very unlikely** under H_0 —strong evidence against it.

Confidence Intervals $\leftrightarrow$ Hypothesis Testing

The Duality:

Testing $H_0 : \mu = \mu_0$ at level α is equivalent to checking whether μ_0 falls inside the $(1 - \alpha) \times 100\%$ confidence interval.

- $\mu_0 \in \text{CI} \Rightarrow$ Fail to reject H_0
- $\mu_0 \notin \text{CI} \Rightarrow$ Reject H_0



Confidence intervals and hypothesis tests are **two sides of the same coin**. Every CI implicitly tests infinitely many hypotheses: any value *inside* the interval cannot be rejected, and any value *outside* is rejected.

Z-Test Procedure: Known Population Variance

Setting: Random sample of n observations from $N(\mu, \sigma^2)$ with σ^2 **known**.

Test Statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1) \quad \text{under } H_0$$

Rejection Rules:

Test Type	Hypotheses	Reject H_0 if
Upper-tail	$H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$	$Z > z_\alpha$
Lower-tail	$H_0 : \mu \geq \mu_0$ vs. $H_1 : \mu < \mu_0$	$Z < -z_\alpha$
Two-tailed	$H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$	$ Z > z_{\alpha/2}$

Alternatively, reject H_0 if $\bar{x} > \mu_0 + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$ (upper-tail case).

Example 5.1: Are Phone Bills Over \$52?

Problem: A phone industry manager thinks that customer monthly cell phone bills have increased and now average over \$52 per month. The company wishes to test this claim.

Given:

- Sample size: $n = 64$
- Sample mean: $\bar{x} = 53.1$
- Population standard deviation: $\sigma = 10$ (known)
- Significance level: $\alpha = 0.10$

Example 5.1: Are Phone Bills Over \$52?

Problem: A phone industry manager thinks that customer monthly cell phone bills have increased and now average over \$52 per month. The company wishes to test this claim.

Given:

- Sample size: $n = 64$
- Sample mean: $\bar{x} = 53.1$
- Population standard deviation: $\sigma = 10$ (known)
- Significance level: $\alpha = 0.10$

Step 1: State the hypotheses.

$H_0 : \mu \leq 52$ (the average is not over \$52 per month)

$H_1 : \mu > 52$ (the average exceeds \$52—manager's claim)

This is a **one-sided, upper-tail test** since the alternative focuses on the upper tail above 52.

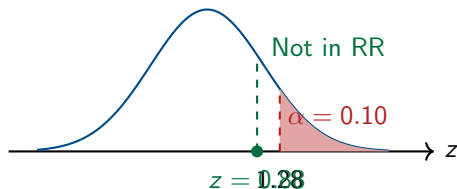
Example 5.1: Critical Value Approach

Step 2: Compute the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{53.1 - 52}{10/\sqrt{64}} = \frac{1.1}{1.25} = 0.88$$

Step 3: Determine the critical value.

For $\alpha = 0.10$ (right-tailed): $z_\alpha = z_{0.10} = 1.28$



Step 4: Decision. Since $z = 0.88 < 1.28 = z_\alpha$, we **fail to reject H_0** .

Example 5.1: P -Value and CI Approaches

P -Value Approach:

$$p\text{-value} = \mathbb{P}(Z > 0.88) = 1 - \Phi(0.88) = 1 - 0.8106 = 0.1894$$

Since $p\text{-value} = 0.1894 > 0.10 = \alpha$, we **fail to reject** H_0 .

Example 5.1: P -Value and CI Approaches

P -Value Approach:

$$p\text{-value} = \mathbb{P}(Z > 0.88) = 1 - \Phi(0.88) = 1 - 0.8106 = 0.1894$$

Since $p\text{-value} = 0.1894 > 0.10 = \alpha$, we **fail to reject** H_0 .

Confidence Interval Approach:

A 90% confidence interval for μ :

$$\bar{x} \pm z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 53.1 \pm 1.645 \times \frac{10}{\sqrt{64}} = 53.1 \pm 2.056 = (51.04, 55.16)$$

Since $\mu_0 = 52$ falls **within** this interval, we cannot reject H_0 .

All three approaches give the **same conclusion**: there is insufficient statistical evidence at the 10% level to support the manager's claim that average monthly phone bills exceed \$52.

Example 5.2: Is the Mean Number of TV Sets Equal to 3?

Problem: Test the claim that the true mean number of TV sets in US homes is equal to 3.

Given: $\sigma = 0.8$, $n = 100$, $\bar{x} = 2.84$, $\alpha = 0.05$

Example 5.2: Is the Mean Number of TV Sets Equal to 3?

Problem: Test the claim that the true mean number of TV sets in US homes is equal to 3.

Given: $\sigma = 0.8$, $n = 100$, $\bar{x} = 2.84$, $\alpha = 0.05$

Hypotheses (two-tailed test):

$$H_0 : \mu = 3 \quad (\text{mean number of TV sets is 3})$$

$$H_1 : \mu \neq 3 \quad (\text{mean number of TV sets is not 3})$$

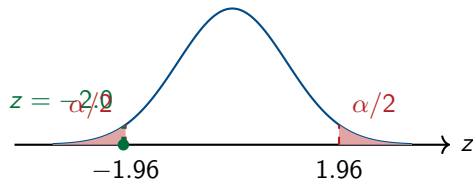
Test statistic:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.84 - 3}{0.8/\sqrt{100}} = \frac{-0.16}{0.08} = -2.0$$

Critical values: $z_{\alpha/2} = z_{0.025} = \pm 1.96$

Since $|z| = 2.0 > 1.96$, we **reject** H_0 .

Example 5.2: All Four Approaches Agree



- ① **Critical value:** $z = -2.0 < -1.96$ ✓ Reject
- ② **P-value:** $2 \times \mathbb{P}(Z < -2.0) = 2 \times 0.0228 = 0.0456 < 0.05$ ✓ Reject
- ③ **95% CI:** $2.84 \pm 1.96 \times 0.08 = (2.683, 2.997)$. Since $3 \notin (2.683, 2.997)$ ✓ Reject
- ④ **Critical boundaries for $\bar{x}$:** $3 \pm 1.96 \times 0.08 = (2.843, 3.157)$. Since $2.84 < 2.843$ ✓ Reject

Conclusion: Sufficient evidence at 5% level to reject the claim that the true mean number of TV sets is equal to 3.

Economics Application 1: Did a Policy Change the Mean?

Scenario: Before a minimum wage increase, the average weekly earnings in a sector were \$480. After the increase, a random sample of $n = 100$ workers shows $\bar{x} = 495$ with $\sigma = 60$ (known from administrative data).

Question: Has the average weekly earnings changed?

Economics Application 1: Did a Policy Change the Mean?

Scenario: Before a minimum wage increase, the average weekly earnings in a sector were \$480. After the increase, a random sample of $n = 100$ workers shows $\bar{x} = 495$ with $\sigma = 60$ (known from administrative data).

Question: Has the average weekly earnings changed?

$$H_0 : \mu = 480 \quad (\text{no change in average earnings})$$

$$H_1 : \mu \neq 480 \quad (\text{earnings have changed})$$

$$z = \frac{495 - 480}{60/\sqrt{100}} = \frac{15}{6} = 2.5$$

At $\alpha = 0.05$: $|z| = 2.5 > 1.96 \Rightarrow$ **Reject** H_0 . p -value
 $= 2 \times \mathbb{P}(Z > 2.5) = 2 \times 0.0062 = 0.0124$

There is statistically significant evidence that average weekly earnings have changed after the minimum wage increase. Note: this does not prove the policy *caused* the change—for causal inference, we need additional econometric methods.

Economics Application 2: Testing Market Efficiency

Scenario: The efficient market hypothesis (EMH) predicts that average abnormal stock returns should be zero. An analyst examines $n = 250$ trading days and finds mean abnormal return $\bar{x} = 0.12\%$ with $\sigma = 1.5\%$.

$H_0 : \mu = 0$ (market is efficient, no abnormal returns)

$H_1 : \mu \neq 0$ (market is inefficient)

Economics Application 2: Testing Market Efficiency

Scenario: The efficient market hypothesis (EMH) predicts that average abnormal stock returns should be zero. An analyst examines $n = 250$ trading days and finds mean abnormal return $\bar{x} = 0.12\%$ with $\sigma = 1.5\%$.

$H_0 : \mu = 0$ (market is efficient, no abnormal returns)

$H_1 : \mu \neq 0$ (market is inefficient)

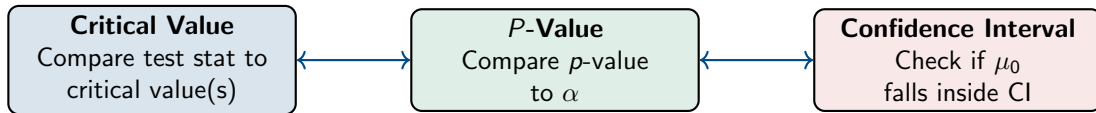
$$z = \frac{0.12 - 0}{1.5/\sqrt{250}} = \frac{0.12}{0.0949} = 1.26$$

At $\alpha = 0.05$: $|z| = 1.26 < 1.96 \Rightarrow$ **Fail to reject H_0 .**

$p\text{-value} = 2 \times \mathbb{P}(Z > 1.26) = 2 \times 0.1038 = 0.2076$

We fail to reject market efficiency. The observed 0.12% mean abnormal return is *not* statistically significant—it could easily have arisen by chance even if the true mean is zero.

Three Equivalent Approaches to Hypothesis Testing



All three approaches always give the same conclusion!

Choose whichever is most convenient or informative for your problem.

- The **critical value approach** is the most “mechanical”—compare numbers
- The **p -value** is the most informative—it tells you *how much* evidence against H_0
- The **CI approach** gives you a range of plausible parameter values

Practice Problem 1: Inflation Target

Problem: The Central Bank's inflation target is 5%. An economist collects a random sample of $n = 36$ monthly price indices and finds a mean inflation rate of $\bar{x} = 5.8\%$. The population standard deviation is known to be $\sigma = 2.4\%$. Test whether inflation exceeds the target at $\alpha = 0.05$.

Practice Problem 1: Inflation Target

Problem: The Central Bank's inflation target is 5%. An economist collects a random sample of $n = 36$ monthly price indices and finds a mean inflation rate of $\bar{x} = 5.8\%$. The population standard deviation is known to be $\sigma = 2.4\%$. Test whether inflation exceeds the target at $\alpha = 0.05$.

Solution:

$$H_0 : \mu \leq 5 \quad \text{vs.} \quad H_1 : \mu > 5 \quad (\text{upper-tail test})$$

$$z = \frac{5.8 - 5}{2.4/\sqrt{36}} = \frac{0.8}{0.4} = 2.0$$

Critical value: $z_{0.05} = 1.645$

Since $z = 2.0 > 1.645$, we **reject** H_0 .

$p\text{-value} = \mathbb{P}(Z > 2.0) = 0.0228 < 0.05 \quad \checkmark$

Conclusion: There is sufficient evidence at the 5% level to conclude that inflation exceeds the 5% target.

Practice Problem 2: Average Commute Time

Problem: A city planner claims that the average commute time in Ankara is 40 minutes. A random sample of $n = 49$ commuters yields $\bar{x} = 43$ minutes. Assume $\sigma = 14$ minutes. Test at $\alpha = 0.01$.

Practice Problem 2: Average Commute Time

Problem: A city planner claims that the average commute time in Ankara is 40 minutes. A random sample of $n = 49$ commuters yields $\bar{x} = 43$ minutes. Assume $\sigma = 14$ minutes. Test at $\alpha = 0.01$.

Solution:

$$H_0 : \mu = 40 \quad \text{vs.} \quad H_1 : \mu \neq 40 \quad (\text{two-tailed test})$$

$$z = \frac{43 - 40}{14/\sqrt{49}} = \frac{3}{2} = 1.5$$

Critical values: $\pm z_{0.005} = \pm 2.576$

Since $|z| = 1.5 < 2.576$, we **fail to reject** H_0 .

$$p\text{-value} = 2 \times \mathbb{P}(Z > 1.5) = 2 \times 0.0668 = 0.1336 > 0.01$$

99% CI: $43 \pm 2.576 \times 2 = (37.85, 48.15)$. Since $40 \in (37.85, 48.15)$ ✓

Conclusion: Insufficient evidence at 1% level to reject the claim that average commute time is 40 minutes.

Practice Problem 3: Effect of Training on Productivity

Problem: A factory's average daily output was 200 units. After a training program, a sample of $n = 81$ days shows $\bar{x} = 206$ units. Assume $\sigma = 27$ units. Has productivity increased? Test at $\alpha = 0.05$.

Practice Problem 3: Effect of Training on Productivity

Problem: A factory's average daily output was 200 units. After a training program, a sample of $n = 81$ days shows $\bar{x} = 206$ units. Assume $\sigma = 27$ units. Has productivity increased? Test at $\alpha = 0.05$.

Solution:

$$H_0 : \mu \leq 200 \quad \text{vs.} \quad H_1 : \mu > 200 \quad (\text{upper-tail test})$$

$$z = \frac{206 - 200}{27/\sqrt{81}} = \frac{6}{3} = 2.0$$

Critical value: $z_{0.05} = 1.645$

Since $z = 2.0 > 1.645$, we **reject** H_0 .

$p\text{-value} = \mathbb{P}(Z > 2.0) = 0.0228 < 0.05 \quad \checkmark$

Conclusion: There is sufficient evidence at the 5% level to conclude that productivity has increased after the training program.

Practice Problem 4: Testing a Wage Claim

Problem: A labor union claims that the average hourly wage in a sector is less than 120 TL. A random sample of $n = 100$ workers gives $\bar{x} = 117$ TL with $\sigma = 20$ TL. Test at $\alpha = 0.05$.

Practice Problem 4: Testing a Wage Claim

Problem: A labor union claims that the average hourly wage in a sector is less than 120 TL. A random sample of $n = 100$ workers gives $\bar{x} = 117$ TL with $\sigma = 20$ TL. Test at $\alpha = 0.05$.

Solution:

$$H_0 : \mu \geq 120 \quad \text{vs.} \quad H_1 : \mu < 120 \quad (\text{lower-tail test})$$

$$z = \frac{117 - 120}{20/\sqrt{100}} = \frac{-3}{2} = -1.5$$

Critical value: $-z_{0.05} = -1.645$

Since $z = -1.5 > -1.645$ (not in the rejection region), we **fail to reject** H_0 .

$p\text{-value} = \mathbb{P}(Z < -1.5) = 0.0668 > 0.05$

Conclusion: Insufficient evidence at the 5% level to support the union's claim that the average hourly wage is below 120 TL.

Notice: $z = -1.5$ is close to the critical value -1.645 . At $\alpha = 0.10$, we *would* reject H_0 . The conclusion depends on the chosen α

Hypothesis Testing: The Complete Recipe

- 1 **State the hypotheses:** H_0 and H_1
 - Determine if one-sided or two-sided
- 2 **Choose the significance level** α (before looking at the data!)
- 3 **Compute the test statistic:**

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

- 4 **Determine the critical value(s)** from the Z-table
- 5 **Make the decision:**
 - Reject H_0 if test statistic falls in the rejection region
 - Or: reject H_0 if $p\text{-value} < \alpha$
 - Or: reject H_0 if $\mu_0 \notin \text{CI}$
- 6 **State your conclusion in context**
 - Remember the two deadly sins!

Common Mistakes to Avoid

Mistake 1: Confusing Statistical and Practical Significance

A result can be **statistically significant** but **practically unimportant**. With a large enough sample, even a tiny difference from μ_0 can be “significant.” Always consider the **magnitude** of the effect.

Mistake 2: Interpreting the P -Value Incorrectly

The p -value is **NOT** the probability that H_0 is true. It is the probability of observing data this extreme *if H_0 were true*.

Mistake 3: Using α After Seeing the Data

The significance level α must be chosen **before** the test is conducted. Adjusting α after seeing the results is a form of data manipulation.

THE TWO DEADLY SINS

1. Never say “Accept H_0 .”

Say: “Fail to reject H_0 ” or “Insufficient evidence to reject H_0 .”

2. Never include H_a in your concluding statement.

Your conclusion is always framed in terms of H_0 .

Correct: “We fail to reject $H_0 : \mu = 52$ at the 10% significance level.”

Wrong: “We accept that $\mu = 52$.”

Wrong: “We conclude that $\mu > 52$.” (This mentions H_a !)

Key Takeaways from Week 9

- 1 Hypothesis testing answers “**Is this value plausible?**” while estimation answers “What is the value?”
- 2 The logic is “**innocent until proven guilty**”: H_0 is maintained unless the evidence overwhelmingly contradicts it
- 3 The four components: H_0 , H_1 , test statistic, rejection region
- 4 **Type I error** (α): rejecting a true H_0 ; **Type II error** (β): failing to reject a false H_0
- 5 Three equivalent approaches: critical value, p -value, confidence interval
- 6 CI and hypothesis testing are **two sides of the same coin**
- 7 **Never say “Accept H_0 ” and never include H_a in your conclusion**

Coming Up Next Week

Hypothesis testing with **unknown variance** (t -test), testing for proportions, and testing for population variances (χ^2 -test).

Further Reading

- Wackerly, Mendenhall & Scheaffer, Ch. 10: Hypothesis Testing
- Newbold, Carlson & Thorne, Ch. 9: Hypothesis Testing (single population)
- Hogg, Tanis & Zimmerman, Ch. 6: Tests of Statistical Hypotheses
- Devore & Berk, Ch. 8: Tests of Hypotheses Based on a Single Sample