ECO 391 Economics and Business Statistics
Lecture 5: Hypothesis Testing

Xiaozhou Ding

February 13, 2019

Overview

Introduction to Hypothesis Testing

Introduction to Hypothesis Testing

- Hypothesis tests resolve conflicts between two competing opinions (hypotheses).
- In a hypothesis test, define
  - H_0, the **null hypothesis**, the presumed default state of nature or status quo.
  - H_A, the **alternative hypothesis**, a contradiction of the default state of nature or status quo.

Introduction to Hypothesis Testing

- In statistics we use sample information to make inferences regarding the unknown population parameters of interest.
- We conduct hypothesis tests to determine if sample evidence contradicts H_0.
- On the basis of sample information, we either
  - "Reject the null hypothesis".
    - Sample evidence is inconsistent with H_0.
  - "Fail to reject the null hypothesis".
    - Sample evidence is not inconsistent with H_0.
    - We do not have enough evidence to "accept" H_0.

Defining the Null Hypothesis and Alternative Hypothesis

- General guidelines:
    - Null hypothesis H_0 states the status quo.
    - Alternative hypothesis H_A, states whatever we wish to establish (i.e., contests the status quo).
- Use the following signs in hypothesis tests

H_0	$=$	$\geq$	$\leq$	Status quo
H_A	$\neq$	$<$	$>$	Contradict H_0

- Note that H_0 always contains the "equality".

One-Tailed versus Two-Tailed Hypothesis Tests

- Two-Tailed Test
  - Reject H_0 on either side of the hypothesized value of the population parameter.
  - For example: $H_0 : \mu = \mu_0$ versus $H_A : \mu \neq \mu_0$.
    The "$\neq$" symbol in H_A indicates that both tail areas of the distribution will be used to make the decision regarding the rejection of H_0.
- One-Tailed Test
  - Reject H_0 only on one side of the hypothesized value of the population parameter.
  - For example:
    - ★ $H_0 : \mu \leq \mu_0$ versus $H_A : \mu > \mu_0$ (right-tail test)
    - ★ $H_0 : \mu \geq \mu_0$ versus $H_A : \mu < \mu_0$ (left-tail test)

    Note that the inequality in H_A determines which tail area will be used to make the decision regarding the rejection of H_0.

Three Steps to Formulate Hypotheses

1. Identify the relevant population parameter of interest (e.g., μ).
2. Determine whether it is a one- or a two-tailed test.
3. Include some form of the equality sign in H_0 and use H_A to establish a claim.

Intuition

- We can never really prove that the true population parameter is equal to a specified value, but we can prove that the parameter is NOT equal to some specified value.
- Thus, we apply a principle of "proving by a statistical contradiction".
  - The proposition in H_0 is a sort of "default case" and we assume it to be "valid" unless sample evidence throws serious doubts on it.
  - H_0 should be what you'd like to disprove.
  - H_A should have what you are trying to prove - the statement that has a burden of proof.

Useful Tips

- Generally, we use the alternative hypotheses as a vehicle to establish something or contest the status quo for which a corrective action may be required.
- Propositions with $=, \geq, \leq$ always go to H_0.
- Propositions with $\neq, <, >$ always go to H_0.

Example

A trade group predicts that back-to-school spending will average $606.40 per family this year. A different economic model is needed if the prediction is wrong.

1. Parameter of interest is μ since we are interested in the average back-to-school spending.

2. Since we want to determine if the population mean differs from $606.4 (i.e,), it is a two-tail test.

3. $H_0 : \mu = 606.4$, $H_A : \mu \neq 606.4$.

Example

A television research analyst wishes to test a claim that more than 30% of the households will tune in for a TV episode. Specify the null and the alternative hypothesis to test the claim.

Example

A television research analyst wishes to test a claim that more than 30% of the households will tune in for a TV episode. Specify the null and the alternative hypothesis to test the claim.

This is a one-tailed test. $H_0 : p \leq 0.3$, $H_A : p > 0.3$.

Example

Suppose that a government inspector is concerned about whether a particular cereal producer is putting an average of **at least** 32 ounces of cereal into each box as claimed. What are your null and alternative hypotheses?

Example

Suppose that a government inspector is concerned about whether a particular cereal producer is putting an average of **at least** 32 ounces of cereal into each box as claimed. What are your null and alternative hypotheses?

$H_0 : \mu \geq 32$, $H_A : \mu < 32$. One-tailed test.

Type I and Type II Errors

- Type I Error: Committed when we reject H_0 when H_0 is actually true.
  - Occurs with probability α.
- Type II Error: Committed when we do not reject H_0 and H_0 is actually false.
  - Occurs with probability β.
  - Power of the test $= 1 - \beta$. It is defined as the probability of rejecting the null hypothesis when the null hypothesis is false.

For a given sample size n, a decrease in α will increase β and vice versa. Both α and β decrease as n increases.

Type I and Type II Errors

- This table illustrates the decisions that may be made when hypothesis testing:

Decision	Null hypothesis is true	Null hypothesis is false
Reject the null hypothesis	Type I error	Correct decision
Do not reject the null hypothesis	Correct decision	Type II error

- Correct Decisions:
  - Reject H_0 when H_0 is false.
  - Do not reject H_0 when H_0 is true.
- Incorrect Decisions:
  - Reject H_0 when H_0 is true (Type I).
  - Do not reject H_0 when H_0 is false (Type II).

Example

Consider the following competing hypotheses that relate to the court of law.

- H_0: An accused person is innocent.
- H_A: An accused person is guilty.

Consequences of Type I and Type II errors:

- Type I error: Reject H_0 and conclude that the accused is guilty when in reality, she is innocent.
- Type II error: Do not reject the H_0 and conclude that the accused is innocent when in reality, she is guilty.

Hypothesis Test of the Population Mean When σ Is Known

Introduction

- Hypothesis testing enables us to determine whether the sample evidence is inconsistent with what is hypothesized under the null hypothesis (H_0).
- Basic principle: First assume that H_0 is true and then determine if sample evidence contradicts this assumption.

The *p*-value Approach

- The value of the test statistic for the hypothesis test of the population mean μ when the population standard deviation σ is known is computed as

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

  where μ_0 is the hypothesized mean value.

- z is called our test statistic or our "observed z-score".
- This formula is valid only if $\overline{X}$ (approximately) follows a normal distribution.
- *p*-value: the likelihood of obtaining a sample mean that is at least as extreme as the one derived from the given sample, under the assumption that the null hypothesis is true.

Note: it's equivalent to "If the defendant is innocent, what is the chance that we'd observe such extreme criminal evidence?"
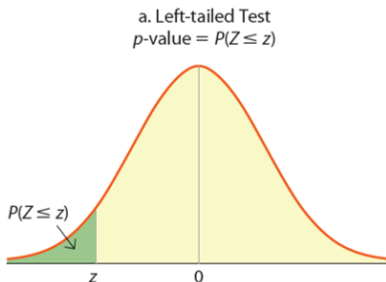
The *p*-value Approach

- Under the assumption that $\mu = \mu_0$, the *p*-value is the likelihood of observing a sample mean that is at least as extreme as the one derived from the given sample.
- The calculation of the *p*-value depends on the specification of the alternative hypothesis:

Alternative Hypothesis	*p*-value
$H_A: \mu > \mu_0$	Right-tail probability: $P(Z \geq z)$
$H_A: \mu < \mu_0$	Left-tail probability: $P(Z \leq z)$
$H_A: \mu \neq \mu_0$	Two-tail probability: $2P(Z \geq z)$ if $z > 0$ or $2P(Z \leq z)$ if $z < 0$

- Decision rule: Reject H_0 if *p*-value $< \alpha$.
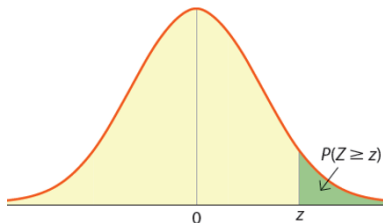
The *p*-value Approach

Determining the *p*-value depending on the specification of the competing hypotheses.

a. Left-tailed Test
p-value = $P(Z \leq z)$

$P(Z \leq z)$

The *p*-value Approach

Determining the *p*-value depending on the specification of the competing hypotheses.

b. Right-tailed Test
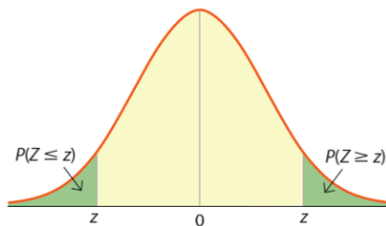*p*-value $= P(Z \geq z)$

$P(Z \geq z)$

0   z

The *p*-value Approach

Determining the *p*-value depending on the specification of the competing hypotheses.

c. Two-tailed Test
If $z < 0$, then *p*-value $= 2P(Z \leq z)$
If $z > 0$, then *p*-value $= 2P(Z \geq z)$

$P(Z \leq z)$

$P(Z \geq z)$

z 0 z

Four Step Procedure Using The *p*-value Approach

1. Specify the null and the alternative hypotheses.
2. Specify the test statistic and compute its value.
3. Calculate the *p*-value.
4. Set significance level, state the conclusion and interpret the results.

Set the significance level, α, the probability of making a Type I error to be small - 0.01, 0.05, or 0.10. Compare the *p*-value to α. If the *p*-value is less than (or equal to) α, reject the null hypothesis in favor of the alternative hypothesis.

Intuition: If we set $\alpha = 0.05$ means we allow 5% chance of Type I error of rejecting the defendant being innocent but actually is innocent. However, if $p < \alpha$, meaning the sample tells us we will be less likely to make the mistake, so reject H_0 and support H_A.

Example

Suppose a sociologist wants to establish that the mean retirement age is greater than 67. It is assumed that retirement age is normally distributed with a known population standard deviation of 9 years. Let a random sample of 25 retirees produce an average retirement age of 71. This sample evidence casts doubts on the validity of the null hypothesis.

Example

$n = 25$, $\bar{x} = 71$, $\sigma = 9$.

1. Step 1: State the hypotheses:

$$\begin{cases} H_0: & \mu \le 67 \\ H_A, & \mu > 67. \end{cases}$$

Thus, $\mu_0 = 67$.

2. Step 2: Given that the population is normally distributed with a known standard deviation, $\sigma = 9$, we compute the value of the test statistic as

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{71 - 67}{9/\sqrt{25}} = 2.22$$
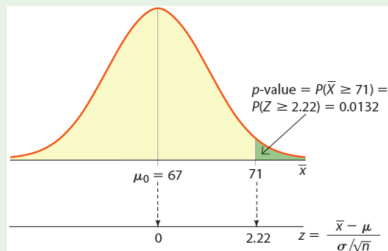
Example

Unstandardized Normal Distribution: $\bar{x} = 71, \mu_0 = 67$. Standardized Normal Distribution: $z = 2.22, \mu = 0$.

3. Step 3: Now compute the *p*-value. Note that since $H_A : \mu > 67$, this is a right-tail test. Thus,

$$P(\overline{X} \geq 71) = P(Z \geq 2.22) = 1 - 0.9868 = 0.0132$$

p-value is 0.0132 or 1.32%.

Typically, before implementing a hypothesis test, we choose a value for α (level of significance) = 0.01, 0.05, or 0.1 and reject H_0 when the *p*-value < α.

Let's say, before conducting the study, we chose $\alpha = 0.05$.

Example

④ Step 4: Since p-value $= 0.0132 < \alpha = 0.05$, at the 5% level of significance, we reject H_0 and conclude that the sample data support the alternative claim that $\mu > 67$.

Interpret Your Results

- Case I: When we reject H_0.
  When we reject H_0 at the $\alpha\%$ level of significance, then the proposition in H_A is referred to as statistically significant at the $\alpha\%$ level.
  If we reject H_0 at 5% level of significance, then will we also reject H_0
    - at 1%? NO (we are 95% sure but not 99% sure).
    - at 10%? YES (If we are 95% sure we are also 90% sure.)
- Case II: When we fail to reject H_0.
  We do not "prove H_0", we just **fail to reject it**.
  When we fail to reject H_0 at the $\alpha\%$ level of significance, H_A is referred to as statistically insignificant at $\alpha\%$ level.

The Critical Value Approach

- Rejection region: a region of values such that if the test statistic falls into this region, then we reject H_0.
  - The location of this region is determined by H_A.
- Critical value: a point that separates the rejection region from the non-rejection region.
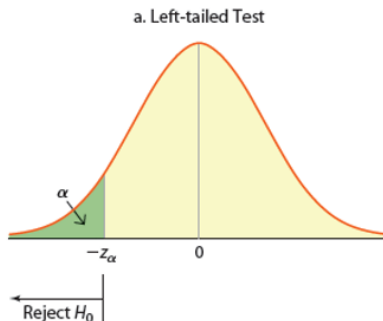
The Critical Value Approach

- The critical value approach specifies a region such that if the value of the test statistic falls into the region, the null hypothesis is rejected.
- The critical value depends on the alternative hypothesis.

Alternative Hypothesis	Critical Value
$H_A: \mu > \mu_0$	Right-tailed critical value is z_α, where $P(Z \geq z_\alpha) = \alpha$.
$H_A: \mu < \mu_0$	Left-tailed critical value is $-z_\alpha$, where $P(Z \leq -z_\alpha) = \alpha$.
$H_A: \mu \neq \mu_0$	Two-tailed critical values $-z_{\alpha/2}$ and $z_{\alpha/2}$, where $P(Z \geq z_{\alpha/2}) = \alpha/2$.

- Decision Rule: Reject H_0 if:
    - $z > z_\alpha$ for a right-tailed test
    - $z < -z_\alpha$ for a left-tailed test
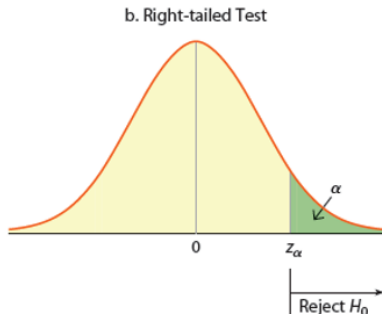   - $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$ for a two-tailed test.

The Critical Value Approach

Determining the critical value(s) depending on the specification of the competing hypotheses.

a. Left-tailed Test

The Critical Value Approach

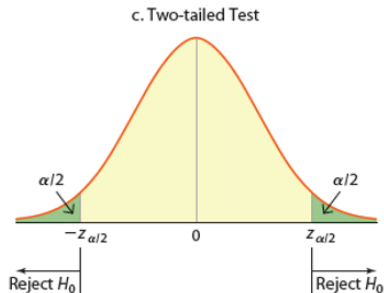Determining the critical value(s) depending on the specification of the competing hypotheses.

b. Right-tailed Test

The Critical Value Approach

Determining the critical value(s) depending on the specification of the competing hypotheses.

c. Two-tailed Test

$\alpha/2$

$\alpha/2$

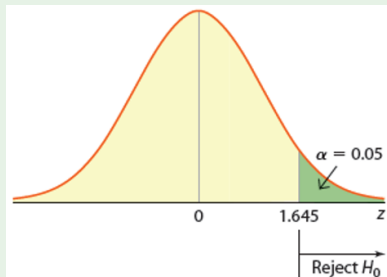$-z_{\alpha/2}$

0

$z_{\alpha/2}$

Reject H_0

Reject H_0

Four Step Procedure Using the Critical Value Approach

1. Specify the null and the alternative hypotheses.
2. Specify the test statistic and compute its value.
3. Set significance level and find the critical value or values.
4. State the conclusion and interpret the results.

> **Example**
>
> 1. $H_0 : \mu \le 67$, $H_A : \mu > 67$.
> 2. From previous example, $z = 2.22$.
> 3. Based on H_A, this is a right-tail test and for $\alpha = 0.05$, the critical value is $z_\alpha = z_{0.05} = 1.645$.
>
>
>
> 4. Reject H_0 if $z > 1.645$.
>   Since $z = 2.22 > z_\alpha = 1.645$, the test statistic falls in the rejection region. Therefore, we reject H_0 and conclude that the sample data support the alternative claim $\mu > 67$.

This conclusion is the same as that from the *p*-value approach.

Linking the Two Methods

- If z falls in the rejection region, then the p-value must be less than α.
- If z does not fall in the rejection region, then the p-value must be greater than α.
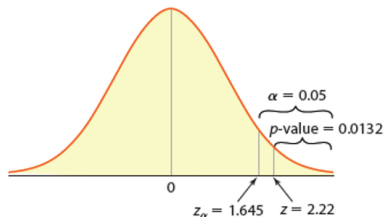
Confidence Intervals and Two-Tailed Hypothesis Tests

- Given the significance level α, we can use the sample data to construct a $100(1 - \alpha)\%$ confidence interval for the population mean μ.
- Decision rule
  - Reject H_0 if the confidence interval does not contain the value of the hypothesized mean μ_0.
  - Do not reject H_0 if the confidence interval does contain the value of the hypothesized mean μ_0.

Implementing a Two-Tailed Test Using a Confidence Interval

- The general specification for a $100(1 - \alpha)\%$ confidence interval of the population mean μ when the population standard deviation σ is known is computed as

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

- Decision rule: Reject H_0 if
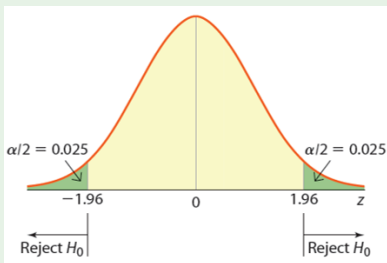
$$\mu_0 < \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

or if

$$\mu_0 > \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Example

A research analyst wishes to determine if average back-to-school spending differs from $606.40. Out of 30 randomly drawn households from a normally distributed population, the standard deviation is $65 and sample mean is $622.85.

1. $H_0 : \mu = 606.4$, $H_A : \mu \neq 606.4$.

2. $z = 1.39$

3. Based on H_A, this is a two-tail test and for $\alpha = 0.05$, the critical value is $z_{\alpha/2} = z_{0.025} = \pm 1.96$.

Example

1. The confidence interval is

$$\left[622.85 - 1.96 \times \frac{65}{\sqrt{30}}, [622.85 + 1.96 \times \frac{65}{\sqrt{30}} \right],$$

$$[599.59, 646.11]$$

2. It includes \$606.40, so fail to reject.

Example

The job placement center at a university claims that the mean starting salaries of college graduates from their institution is no less than \$26,000. (Assume that the population has a $\sigma = \$4,000$.) However, recent graduates have refuted this claim saying that starting salaries are much lower. The students take a random sample of 100 recent graduates and find that the average starting salary is \$24,000. Test the students' claim at the 1% level of significance.

- Step 1: Determine your null and alternative hypotheses

$$\begin{cases} H_0 : \mu \geq 26,000 \\ H_A : \mu < 26,000. \end{cases}$$

- Step 2: Calculate your test statistic (observed)

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{24000 - 26000}{4000 \times \sqrt{100}} = -\frac{2000}{400} = -5.$$

- Step 3: Find the critical value
  This is a left-tailed test. $P(Z < -z_\alpha) = \alpha = 0.01$. From the z-table, we know the critical value $-z_\alpha = -2.33$, so $z_\alpha = 2.33$.
  Decision rule: Reject H_0 if $z < -z_\alpha$
  Based on the critical value, determine your rejection and acceptance region.
- Step 4: Make a decision and state your conclusion
  - The test statistic fall into the rejection area. So we reject the null hypothesis. We are 99% confident that the mean staring salary of college graduate is less than \$26,000.
  - The mean staring salary of college graduate is less than \$26,000 statistically significant at 1% level.

Hypothesis Test of the Population Mean When σ is Unknown

Test Statistic for μ When σ is Unknown

- When the population standard deviation σ is unknown, the test statistic for testing the population mean μ is assumed to follow the t_{df} distribution with $(n-1)$ degrees of freedom (df).

- The value of t_{df} is computed as

$$t_{df} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- Critical values are as follows:
  - For a right-tailed test: $t_{\alpha, n-1}$
  - For a left-tailed test: $-t_{\alpha, n-1}$
  - For a two-tailed test: $-t_{\alpha/2, n-1}$ and $+t_{\alpha/2, n-1}$.

Example: Look up the t-statistic for a right-tailed test at the 5% level when $n = 25$.
$t_{0.05,24} = 1.711$.

Example

A University president wonders if students at her university study less than the national average of 24 hours per week. She randomly select 35 students and asks their average study time per week. From their response she calculates a sample mean of 16.37 hours and a sample standard deviation of 7.22 hours.

- Specify the null and alternative hypothesis
- Calculate the value of the test statistic
- At the 5% significance level, specify the critical values and the decision rule
- What is the conclusion of this test?

- Step 1: determine your null and alternative hypothesis

$$\begin{cases} H_0: & \mu \geq 24 \\ H_A: & \mu < 24. \end{cases}$$

- Step 2: calculate the test statistic
  Since $n = 35$, the degree of freedom equals $35 - 1 = 34$. so the test statistic is

$$t_{34} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{16.37 - 24}{7.22/\sqrt{35}} = -6.25$$

- Step 3: Find the critical value and rejection region
  This is a left-tailed test. The critical value is given by $-t_{\alpha,df}$ where
  $P(T_{df} \geq t_{\alpha,df}) = \alpha$. From the t-table, we know $t_{0.05,34} = 1.691$, so the critical value
  $-t_{0.05,34} = -1.691$.
  Reject null hypothesis if test statistic is less than -1.691 (draw a graph to help you understand)

- Step 4: Conclude and interpret your results
  Since the test statistic fall into the rejection area, we reject the null hypothesis. At the 5% significance level, average study time at the university is less than the national average of 24 hours per week.

Summary

- Establish hypothesis testing (what you want to prove goes into alternatives).
- p-value approach: p is related with H_A, hence with α, the Type I error of wrongly reject something that is true. If p is small (comparing to α), we are less likely to make mistakes by rejecting, so reject.
- Critical-value approach: easier to use and understand in a graphical way.
- Confidence interval approach: easier to use once know the significance level.
- Hypothesis testing when σ is unknown, switch from z-statistic to t-statistic.