Summary---Topic 6: Hypothesis Testing

A statistical hypothesis is a claim about the <u>population parameter</u>
E.g. population mean, population standard deviation, or population proportion, etc.

Hypothesis Test for the Population Mean

- ❑ Critical Value Approach
- ❑ p-Value Approach

Hypothesis Testing Procedure

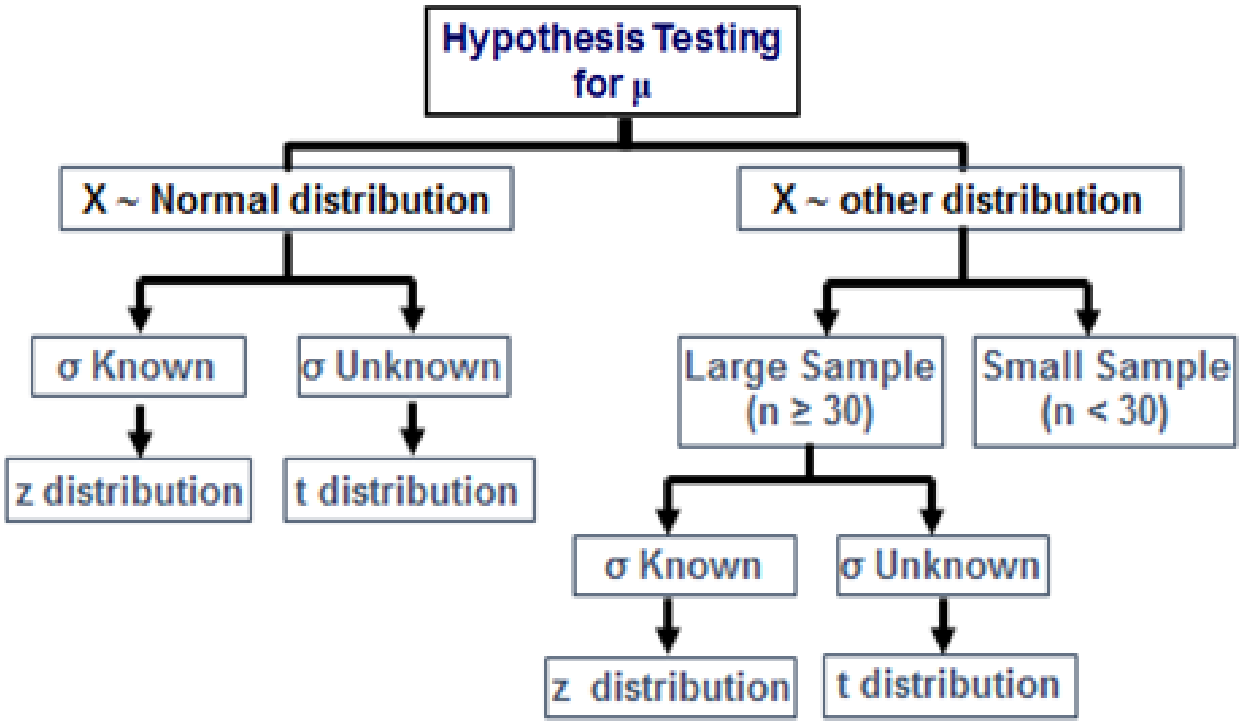
Step 1: Define hypotheses
Step 2: Collect the data and identify the rejection region(s)
Step 3: Compute test statistic
Step 4: Make statistical decision

Critical Value Approach---Step 1: Define Hypotheses

- Always about a population parameter (μ, σ), rather than a sample statistic ($\bar{X}$, s)
- Null hypothesis, H_0: Always contains the "=" sign
- Alternative hypothesis, H_1: Never contains the "=" sign (Mutually exclusive and collectively exhaustive from H_0)
- Always assumed H_0 is true at start (i.e. assume the hypothesis regarding to population parameter is true at start), and then use sample statistics to assess the strength of the evidence against H_0 so as to determine whether H_0 should be rejected or not
- Three sets of hypotheses to be tested

Two-tail test	Lower-tail test	Upper-tail test
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$

Critical Value Approach---
Step 2: Collect the Data and Identify the Rejection Region(s)

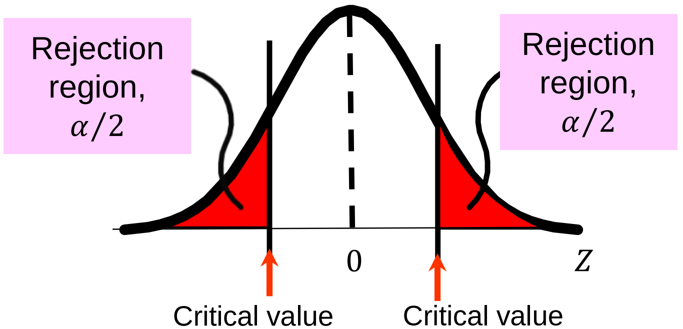


Critical Value Approach---
Step 2: Collect the Data and Identify the Rejection Region(s)

Two-tail test

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

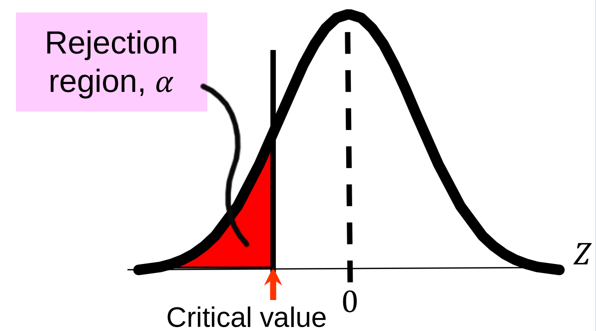


Z **must be** **significantly different from** **0 to reject** H_0

Lower-tail test

$H_0: \mu \geq \mu_0$

$H_1: \mu < \mu_0$

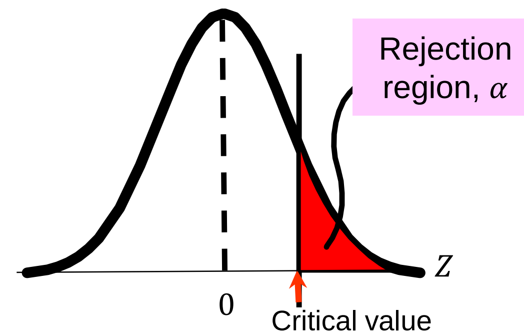


Z **must be** **significantly smaller than** **0 to reject** H_0

Upper-tail test

$H_0: \mu \leq \mu_0$

$H_1: \mu > \mu_0$



Z **must be** **significantly larger than** **0 to reject** H_0

- Choose the **significance level, α,** which specifies the “risk” we are willing to take of rejecting a true H_0; α determines the rejection region and critical value

We reject H_0 if:

Two-tail test	Lower-tail test	Upper-tail test
(Z-distribution) $Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$	(Z-distribution) $Z < -Z_{\alpha}$	(Z-distribution) $Z > Z_{\alpha}$
(t-distribution) $t < -t_{\alpha/2, (n-1)}$ or $t > t_{\alpha/2, (n-1)}$	(t-distribution) $t < -t_{\alpha, (n-1)}$	(t-distribution) $t > t_{\alpha, (n-1)}$

Critical Value Approach---Step 3: Compute Test Statistic

- Test statistic: A statistic that is obtained from the random sample which will be used to test the hypothesis; it is a measure of evidence contained in data that support H_1
- If σ is known, use Z test statistic.

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

- If σ is unknown, use t test statistic, with (n-1) degrees of freedom

$$t = \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$$

- Higher Z-test statistic or t-test statistic, stronger evidence contained in data that support H_1 (i.e. should reject H_0)

Critical Value Approach---Step 4: Make statistical decision

- If the Z or t test statistic value falls in the rejection region, reject H_0. Otherwise, do not reject H_0. (Do not reject $H_0 \neq$ Accept H_0)

- Reject H_0 $\Rightarrow$ There is sufficient evidence that the H_1 is true.
- Do not reject H_0 $\Rightarrow$ There is insufficient evidence that the H_1 is true.

Critical Value Approach--Summary

<table>
<tr><th>Case</th><th>Population Distribution</th><th>Sample Size(n)</th><th>Sampling Distribution of Mean($\bar{X}$)</th><th>σ known, use Z distribution, $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$</th><th>σ unknown, use t distribution, $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$</th></tr>
<tr><td>1</td><td>Normal</td><td>Whatever</td><td>Normal</td>
<td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$
Reject H_0 if:
$Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
Reject H_0 if:
$Z < -Z_{\alpha}$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
Reject H_0 if:
$Z > Z_{\alpha}$</td>
<td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$
Reject H_0 if:
$t < -t_{\alpha/2, (n-1)}$ or $t > t_{\alpha/2, (n-1)}$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
Reject H_0 if:
$t < -t_{\alpha, (n-1)}$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
Reject H_0 if:
$t > t_{\alpha, (n-1)}$</td></tr>
<tr><td>2</td><td>Unknown / not normal</td><td>$n \geq 30$</td><td>By CLT, sample mean $\bar{X}$ is approximately normally distributed</td></tr>
<tr><td>3</td><td>Unknown</td><td>$n < 30$</td><td>Not normal; We need additional assumption. (population distribution is normal, then go to case 1)</td></tr>
</table>

- **Reject H_0 ➔ There is sufficient evidence that the H_1 is true.**
- **Do not reject H_0 ➔ There is insufficient evidence that the H_1 is true.**

p-value approach

Step 1: Define Hypotheses

- Same as the one under the critical value approach

Step 2: Collect the Data and Identify the Rejection Region(s)

- Reject H_0 if <u>p-value < α</u>

Step 3: Compute Test Statistic and p-value

- Obtain the p-value after the computation of test statistic
- p-value: The probability of obtaining a test statistic as extreme or more extreme (≤ or ≥) than the observed sample statistic given H_0 is true
- Smaller p-value (i.e. observed result would be unlikely to occur if H_0 is true), stronger evidence contained in data that support H_1 (i.e. should reject H_0)

Step 4: Make statistical decision

- If p-value < α, then reject H_0; if p-value ≥ α, then do not reject H_0
- Reject H_0 ➔ There is sufficient evidence that the H_1 is true.
- Do not reject H_0 ➔ There is insufficient evidence that the H_1 is true.

p-value Approach--Summary

<table>
<tr><th>Case</th><th>Population Distribution</th><th>Sample Size(n)</th><th>Sampling Distribution of Mean($\bar{X}$)</th><th>σ known, use Z distribution, $z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$</th><th>σ unknown, use t distribution, $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$</th></tr>
<tr><td>1</td><td>Normal</td><td>Whatever</td><td>Normal</td><td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$

p-value= $\mathrm{P}(Z \leq -|z|) + \mathrm{P}(Z \geq |z|)$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
p-value= $\mathrm{P}(Z \leq z)$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
p-value= $\mathrm{P}(Z \geq z)$</td><td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$

p-value= $\mathrm{P}(t_{n-1} \leq -|t|) + \mathrm{P}(t_{n-1} \geq |t|)$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
p-value= $\mathrm{P}(t_{n-1} \leq t)$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
p-value= $\mathrm{P}(t_{n-1} \geq t)$</td></tr>
<tr><td>2</td><td>Unknown / not normal</td><td>n ≥ 30</td><td>By CLT, sample mean $\bar{X}$ is approximately normally distributed</td></tr>
<tr><td>3</td><td>Unknown</td><td>n < 30</td><td>Not normal;
We need additional assumption. (population distribution is normal, then go to case 1)</td></tr>
</table>

- **If p-value < α➔Reject H_0 ➔ There is sufficient evidence that the H_1 is true.**
- **if p-value ≥ α➔Do not reject H_0 ➔ There is insufficient evidence that the H_1 is true.**

Errors in Decision Making

- Type I error: Reject H_0 **given** H_0 is true

→ P(Type I error) = level of significance = α

- Type II error: Do not reject H_0 **given** H_0 is false

→ P(Type II error) = β

Decision	The Truth	
	$\boldsymbol{H_0}$ True	$\boldsymbol{H_0}$ False
Do not reject $\boldsymbol{H_0}$	Right decision Confidence (1- α)	Wrong decision Type II Error (β)
Reject $\boldsymbol{H_0}$	Wrong decision Type I Error (α)	Right decision Power (1- β)

- There would be a **tradeoff** between type I error and type II error. When α ↓, β ↑
- To decrease both errors, we need **increase the sample size n**.

Exercises and Solutions

Q1. Do students at your school study more, less, or about the same as at other business schools? Business Week reported that at the top 50 business schools, students studied an average of 14.6 hours. Set up a hypothesis test to try to prove that the mean number of hours studied at your school is different from the 14.6 hour benchmark reported by Business Week.

a) State the null and alternative hypotheses.

b) What is a Type I error for your test?

c) What is a Type II error for your test?

Solution: a) Let μ be the population mean of study hours

$$H_0: \mu = 14.6 \text{ hours}$$
$$H_1: \mu \neq 14.6 \text{ hours}$$

b)

Type I Error occurs if reject the H_0 when it is true

A Type I error is the mistake of **concluding** that *the mean number of hours studied at your school is **different** from the 14.6 hour benchmark reported by Business Week when **in fact** it is not any different.*

c)

Type II Error occurs if do not reject the H_0 when it is false

A Type II error is the mistake of **not concluding** that *the mean number of hours studied at your school is **different** from the 14.6 hour benchmark reported by Business Week when it is **in fact** different.*

Q2. ATMs must be stocked with enough cash to satisfy customers making withdrawals over an entire weekend. But if too much cash is unnecessarily kept in the ATMs, the bank is forgoing the opportunity of investing the money and earning interest. Suppose that at a particular branch the **population mean** amount of money withdrawn from ATMs per customer transaction over the weekend is **$160** with a **population standard deviation of $30**.

a) If a random sample of 36 customer transactions is examined and the sample mean withdrawal is $148, is there evidence to believe that the population average withdrawal is less than $160? (Use a 0.05 level of significance.)

b) Compute the p-value and interpret its meaning.

Solution:

a) $\mathrm{n} = 36, \;\; \bar{X} = 148, \;\; \sigma = 30, \;\; \alpha = 0.05$

Let μ be the population mean of withdrawal

$H_0: \mu \geq 160$

$H_1: \mu < 160$

Critical Value Approach--Summary

<table>
<tr><th>Case</th><th>Population Distribution</th><th>Sample Size(n)</th><th>Sampling Distribution of Mean($\bar{X}$)</th><th>σ known, use Z distribution, $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$</th><th>σ unknown, use t distribution, $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$</th></tr>
<tr><td>1</td><td>Normal</td><td>Whatever</td><td>Normal</td><td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$
Reject H_0 if:
$Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
Reject H_0 if:
$Z < -Z_{\alpha}$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
Reject H_0 if:
$Z > Z_{\alpha}$</td><td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$
Reject H_0 if:
$t < -t_{\alpha/2, (n-1)}$ or $t > t_{\alpha/2, (n-1)}$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
Reject H_0 if:
$t < -t_{\alpha, (n-1)}$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
Reject H_0 if:
$t > t_{\alpha, (n-1)}$</td></tr>
<tr><td>2</td><td>Unknown / not normal</td><td>$n \geq 30$</td><td>By CLT, sample mean $\bar{X}$ is approximately normally distributed</td></tr>
<tr><td>3</td><td>Unknown</td><td>$n < 30$</td><td>Not normal; We need additional assumption. (population distribution is normal, then go to case 1)</td></tr>
</table>

- **Reject H_0 ➔ There is sufficient evidence that the H_1 is true.**
- **Do not reject H_0 ➔ There is insufficient evidence that the H_1 is true.**

most commonly used values of Z_α

$\boldsymbol{\alpha}$	0.1	0.05	0.025	0.01	0.005
$\mathbf{Z}_\alpha$	1.281552	1.644854	1.959964	2.326348	2.575829

NORM.INV(lower tail probability, mean, s.d.)

$\mathbf{Z}_\alpha$ =NORM.INV(1- $\boldsymbol{\alpha}$, 0,1)
=-NORM.INV($\boldsymbol{\alpha}$, 0,1)

Q2. ATMs must be stocked with enough cash to satisfy customers making withdrawals over an entire weekend. But if too much cash is unnecessarily kept in the ATMs, the bank is forgoing the opportunity of investing the money and earning interest. Suppose that at a particular branch the population mean amount of money withdrawn from ATMs per customer transaction over the weekend is \$160 with a population standard deviation of \$30.

a) If a random sample of 36 customer transactions is examined and the sample mean withdrawal is \$148, is there evidence to believe that the population average withdrawal is less than \$160? (Use a 0.05 level of significance.)

b) Compute the p-value and interpret its meaning.

Solution:

a) $\mathrm{n} = 36, \ \bar{X} = 148, \ \sigma = 30, \ \alpha = 0.05$

Population distribution unknown, since n = 36 > 30, by Central Limit Theorem, the sampling distribution of $\bar{X}$ is approximately normal.

Furthermore, since σ is known (=30), Z test can be used (Lower-tail test) .

Then we compute the test statistic $\mathrm{Z} = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{148-160}{30/\sqrt{36}} = -2.4$,

and the critical value with significance level α is $-Z_\alpha = -Z_{0.05} = -1.645$

Since $\mathrm{Z} = -2.4 < -1.645$, we reject H_0 at $\alpha = 0.05$

There is sufficient evidence that the population mean amount of money withdrawn from ATMs per customer transaction is less than \$160.

p-value Approach--Summary

<table>
<tr><th>Case</th><th>Population Distribution</th><th>Sample Size(n)</th><th>Sampling Distribution of Mean($\bar{X}$)</th><th>σ known, use Z distribution, $z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$</th><th>σ unknown, use t distribution, $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$</th></tr>
<tr><td>1</td><td>Normal</td><td>Whatever</td><td>Normal</td>
<td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$

p-value= $\mathrm{P}(\mathrm{Z} \leq -|\mathrm{z}|) + \mathrm{P}(\mathrm{Z} \geq |z|)$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
p-value= $\mathrm{P}(\mathrm{Z} \leq z)$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
p-value= $\mathrm{P}(\mathrm{Z} \geq z)$</td>
<td rowspan="3">Two-tail test
$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$

p-value= $\mathrm{P}(t_{n-1} \leq -|\mathrm{t}|) + \mathrm{P}(t_{n-1} \geq |t|)$

Lower-tail test
$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$
p-value= $\mathrm{P}(t_{n-1} \leq \mathrm{t})$

Upper-tail test
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$
p-value= $\mathrm{P}(t_{n-1} \geq t)$</td></tr>
<tr><td>2</td><td>Unknown / not normal</td><td>n ≥ 30</td><td>By CLT, sample mean $\bar{X}$ is approximately normally distributed</td></tr>
<tr><td>3</td><td>Unknown</td><td>n < 30</td><td>Not normal; We need additional assumption. (population distribution is normal, then go to case 1)</td></tr>
</table>

- **If p-value < α➔Reject H_0 ➔ There is sufficient evidence that the H_1 is true.**
- **if p-value ≥ α➔Do not reject H_0 ➔ There is insufficient evidence that the H_1 is true.**

Q2. ATMs must be stocked with enough cash to satisfy customers making withdrawals over an entire weekend. But if too much cash is unnecessarily kept in the ATMs, the bank is forgoing the opportunity of investing the money and earning interest. Suppose that at a particular branch the population mean amount of money withdrawn from ATMs per customer transaction over the weekend is \$160 with a population standard deviation of \$30.

a) If a random sample of 36 customer transactions is examined and the sample mean withdrawal is \$148, is there evidence to believe that the population average withdrawal is less than \$160? (Use a 0.05 level of significance.)

b) Compute the p-value and interpret its meaning.

Solution:

b) The test statistic $z = -2.4$, p-value$=$ $\mathrm{P}(\mathrm{Z} \le z) = \mathrm{P}(Z \le -2.4) = 0.0082.$

Probability of obtaining a test statistic -2.4 or less is 0.0082, given H_0 is true.