

# *Chapter 10: Hypothesis Testing*

## *STAT 2601 – Business Statistics*

Esam Mahdi

School of Mathematics and Statistics  
Carleton University

February 26, 2026

# Learning Objectives

By the end of this chapter, you will be able to:

- Formulate the null ( $H_0$ ) and alternative ( $H_1$ ) hypotheses.
- Understand the difference between Type I and Type II errors.
- Distinguish between the z-test and t-test and identify when each should be used.
- Conduct a hypothesis test for a population mean ( $\mu$ ) and population proportion ( $p$ ) using:
  - ① The Critical Value approach.
  - ② The  $p$ -value approach.
- Apply the z-test and t-test for testing the hypotheses about a population mean.
- Interpret the results in a business context.

# Formulating Hypotheses

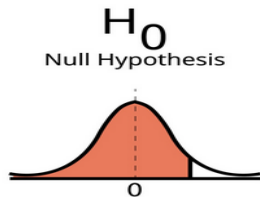
By reading the research problem, create two non-overlapping contrasting situations:

- **The Null Hypothesis ( $H_0$ ):** The statement being tested. It usually represents the "status quo" or no change. It **always** contains the equality sign ( $=$ , or  $\leq$ , or  $\geq$ ).
- **The Alternative Hypothesis ( $H_1$  or  $H_a$ ):** The statement we hope to find evidence for, which is the opposite of the null hypothesis. It never contains an equality sign ( $\neq$ , or  $>$ , or  $<$ ).

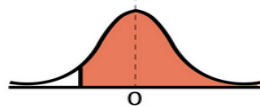
## Examples

- 1 A student claims that a specific study app helps them score higher than their current average of 85.
  - ▶  $H_0 : \mu \leq 85$  (85 denotes the null value  $= \mu_0$ )
  - ▶  $H_1 : \mu > 85$  (Right-tailed test)
- 2 A manager claims a new process reduces assembly time below 15 minutes.
  - ▶  $H_0 : \mu \geq 15$  minutes (15 denotes the null value  $= \mu_0$ )
  - ▶  $H_1 : \mu < 15$  minutes (Left-tailed test)
- 3 A tech firm checks if a new battery's life differs from the standard 8 hours.
  - ▶  $H_0 : \mu = 8$  hours (8 denotes the null value  $= \mu_0$ )
  - ▶  $H_1 : \mu \neq 8$  hours (two-tailed test)

# Non-Overlapping Hypotheses



$H_0: \mu \leq \text{Value}$

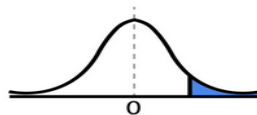


$H_0: \mu \geq \text{Value}$

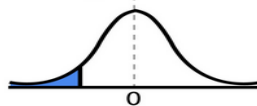


$H_0: \mu = \text{Value}$

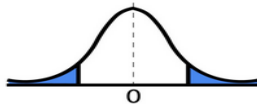
$H_a$   
Alternative Hypothesis



$H_a: \mu > \text{Value}$



$H_a: \mu < \text{Value}$



$H_a: \mu \neq \text{Value}$

# Errors in Hypothesis Testing

In statistics, we never "prove" anything; we only find evidence. This leads to two types of errors:

	$H_0$ <b>True</b>	$H_0$ <b>False</b>
<b>Reject</b> $H_0$	Type I Error ( $\alpha = \text{Significant Level}$ )	Correct Decision = Power ( $1 - \beta$ )
<b>Accept</b> $H_0$	Correct Decision = Confidence ( $1 - \alpha$ )	Type II Error ( $\beta$ )

- **Type I Error:** Rejecting a true null hypothesis when it should not be rejected:

$$\alpha = P(\text{Reject } H_0 | H_0 \text{ true})$$

- **Type II Error:** Accepting the false null hypothesis when it should not be accepted:

$$\beta = P(\text{Accept } H_0 | H_0 \text{ false})$$

# Hypothesis Testing: The Error Trade-off

Unfortunately we cannot minimize both errors ( $\alpha$  and  $\beta$ ) at the same time as decreasing one error necessarily increases the other. Thus, if we need to choose between bad (which is  $\alpha$ ) and worst (which is  $\beta$ ), we have to choose  $\alpha$  and try to avoid making  $\beta$  by never say "Accept  $H_0$ ":

## Correct terminology to avoid making type 2 error

- ✓ **Correct:** We fail to reject the null hypothesis
- ✓ **Correct:** There is insufficient evidence to reject  $H_0$
- ✓ **Correct:** We do not find statistically significant evidence against  $H_0$
- × **Incorrect:** We accept the null hypothesis
- × **Incorrect:** The null hypothesis is true

Note:

- Failing to reject  $H_0$  doesn't prove it's true.
- In practice, we usually set  $\alpha = 0.05$  increasing the sample size in order to increase the power.

# Three Methods for Hypothesis Testing

There are three mathematically equivalent ways to make a decision regarding the Null Hypothesis ( $H_0$ ). At a given significance level  $\alpha$ :

## 1 The Critical Value Method

- ▶ Compare the calculated test statistic to a critical value from the tables.
- ▶ *Rule:* Reject  $H_0$  if the test statistic falls in the "Rejection Region."

## 2 The $p$ -value Method

- ▶ Calculate the probability of obtaining the test statistic as extreme or more extreme than observed, assuming  $H_0$  is true.
- ▶ *Rule:* Reject  $H_0$  if  $p\text{-value} \leq \alpha$ .

## 3 The Confidence Interval (CI) Method

- ▶ Construct a  $(1 - \alpha)100\%$  confidence interval for the parameter.
- ▶ *Rule:* Reject  $H_0$  if the hypothesized value  $\mu_0$  falls **outside** the interval.

## Note

All three methods will always lead to the same statistical conclusion.

# z-Test Statistic for a Population Mean ( $\sigma$ Known)

## Test Statistic

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

where:

- $\bar{x}$ : Sample mean
- $\mu_0$ : Null value (hypothesized population mean under  $H_0$ )
- $\sigma$ : Known population standard deviation
- $n$ : Sample size
- $\sigma / \sqrt{n}$ : Standard error of the sample mean  $\bar{x}$

## Sampling Distribution

Under  $H_0$ , the test statistic (z-score) follows:

$$z \sim N(0, 1)$$

This is the standard normal distribution!



# The Critical Value Method: Step-by-Step

## 1 State the hypotheses: $H_0$ and $H_a$

Right tail  $H_0 : \mu \leq \mu_0$  vs  $H_a : \mu > \mu_0$

Left tail  $H_0 : \mu \geq \mu_0$  vs  $H_a : \mu < \mu_0$

Two tails  $H_0 : \mu = \mu_0$  vs  $H_a : \mu \neq \mu_0$

## 2 Check assumptions: Random sample, normality/large $n$ , known $\sigma$

## 3 Compute test statistic:

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

## 4 Determine critical value(s) (tabular value(s)) based on $\alpha$ and $H_a$

## 5 Make decision:

- ▶ Reject  $H_0$  if test statistic falls in rejection region
- ▶ Fail to reject  $H_0$  otherwise

## 6 State conclusion in context

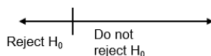
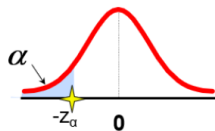
## Rejection Regions

### Lower tail test

Example:

$$H_0: p \geq p_0$$

$$H_A: p < p_0$$

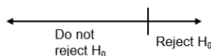
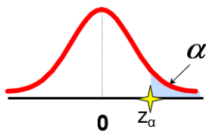


### Upper tail test

Example:

$$H_0: p \leq p_0$$

$$H_A: p > p_0$$

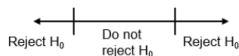
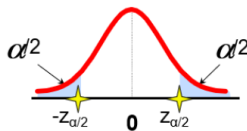


### Two tailed test

Example:

$$H_0: p = p_0$$

$$H_A: p \neq p_0$$



**Rejection Region:** Reject  $H_0$  if:

- Left-tailed test:  $z < -z_\alpha$
- Right-tailed test:  $z > z_\alpha$
- Two-tailed test:  $z > z_{\alpha/2}$  OR  $z < -z_{\alpha/2}$  (or equivalently  $|z| > z_{\alpha/2}$ )

# Critical Values for Common $\alpha$ Levels

Test Type	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.10$
Two-tailed	$\pm 1.96$	$\pm 2.575$	$\pm 1.645$
Right-tailed	1.645	2.33	1.28
Left-tailed	-1.645	-2.33	-1.28

**Table:** Critical z-values for common significance levels

## Finding Critical Values

Use z-table:

- Two-tailed:  $z_{\alpha/2}$  such that  $P(Z > z_{\alpha/2}) = \alpha/2$
- Right-tailed:  $z_{\alpha}$  such that  $P(Z > z_{\alpha}) = \alpha$
- Left-tailed:  $-z_{\alpha}$  (symmetry of normal distribution)

## Example: Lightbulbs Claim (Critical Value Method)

A manufacturer claims their lightbulbs last 1000 hours on average. We test 50 bulbs and find  $\bar{x} = 990$  hours. Assume  $\sigma = 50$  hours. Test at  $\alpha = 0.05$  if the mean lifetime is less than claimed.

### 1 Hypotheses:

- ▶  $H_0 : \mu \geq 1000$
- ▶  $H_a : \mu < 1000$  (left-tailed test)

### 2 Assumptions: Random sample, $n = 50 \geq 30$ , $\sigma$ known

### 3 Test statistic:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{990 - 1000}{50/\sqrt{50}} = \frac{-10}{7.071} = -1.414$$

### 4 Critical value: For left-tailed test with $\alpha = 0.05$ , the critical value is $z_{0.05} = -1.645$

### 5 Decision:

- ▶ Compare:  $z = -1.414$  vs  $z_{0.05} = -1.645$
- ▶ Since  $-1.414 > -1.645$ , test statistic is NOT in rejection region
- ▶ **Fail to reject  $H_0$**

### 6 Conclusion: At $\alpha = 0.05$ , there is insufficient evidence to conclude that the mean lifetime is less than 1000 hours.

# The p-Value Method: Step-by-Step

## 1 State the hypotheses: $H_0$ and $H_a$

Right tail  $H_0 : \mu \leq \mu_0$  vs  $H_a : \mu > \mu_0$

Left tail  $H_0 : \mu \geq \mu_0$  vs  $H_a : \mu < \mu_0$

Two tails  $H_0 : \mu = \mu_0$  vs  $H_a : \mu \neq \mu_0$

## 2 Check assumptions: Random sample, normality/large $n$ , known $\sigma$

## 3 Compute test statistic:

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

## 4 Compute p-value: Probability of obtaining test statistic as extreme or more extreme than observed, assuming $H_0$ is true

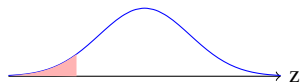
## 5 Make decision:

- ▶ Reject  $H_0$  if p-value  $\leq \alpha$
- ▶ Fail to reject  $H_0$  if p-value  $> \alpha$

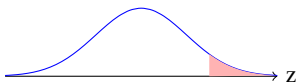
## 6 State conclusion in context

# Calculating p-Values

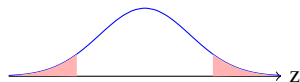
Alternative Hypothesis	p-value Calculation
$H_a : \mu > \mu_0$ (right-tailed)	$P(Z \geq z)$
$H_a : \mu < \mu_0$ (left-tailed)	$P(Z \leq z)$
$H_a : \mu \neq \mu_0$ (two-tailed)	$2 \times P(Z \geq  z )$



Left-tailed:  $p = P(Z < z)$



Right-tailed:  $p = P(Z > z)$



Two-tailed:  $p = 2 \times P(Z > |z|)$

## Interpreting p-values

- **Small p-value:** Unlikely result if  $H_0$  is true  $\Rightarrow$  Evidence against  $H_0$
- **Large p-value:** Likely result if  $H_0$  is true  $\Rightarrow$  Little evidence against  $H_0$
- Common threshold:  $\alpha = 0.05$

## Example: Lightbulbs Claim (p-Value Method)

### Problem Statement (Recall)

Lightbulbs: Claim  $\mu = 1000$ , Sample:  $n = 50$ ,  $\bar{x} = 990$ ,  $\sigma = 50$ ,  $\alpha = 0.05$ ,  $H_a : \mu < 1000$

- 1 **Hypotheses:**  $H_0 : \mu \geq 1000$ ,  $H_a : \mu < 1000$
- 2 **Assumptions:** Satisfied
- 3 **Test statistic:**  $z = -1.414 \approx -1.41$  (calculated earlier)
- 4 **p-value:** For left-tailed test:

$$\text{p-value} = P(Z \leq -1.41)$$

From z-table:  $P(Z \leq -1.411) = 0.0793$  So p-value = 0.0793

- 5 **Decision:**
  - ▶ Compare: p-value = 0.0793 vs  $\alpha = 0.05$
  - ▶ Since  $0.0793 > 0.05$ , p-value  $> \alpha$
  - ▶ **Fail to reject  $H_0$**
- 6 **Conclusion:** At  $\alpha = 0.05$ , there is insufficient evidence to conclude that the mean lifetime is less than 1000 hours.

## Practice Problem (Try This Yourself)

A cereal company claims each box contains 500g of cereal. A consumer group samples 40 boxes and finds  $\bar{x} = 495g$ . Assume  $\sigma = 20g$ . Test at  $\alpha = 0.05$  if the mean weight is different from claimed.

### Instructions:

- 1 State hypotheses
- 2 Check assumptions
- 3 Calculate test statistic
- 4 Use both critical value and p-value methods
- 5 State conclusion

### Solution (Check Your Work)

- $H_0 : \mu = 500$  against  $H_a : \mu \neq 500$
- $z = (495 - 500)/(20/\sqrt{40}) = -1.581 \approx -1.58$
- Critical values:  $\pm 1.96$  (fail to reject)
- p-value:  $2 \times P(Z \geq 1.58) = 2 \times 0.0571 = 0.1142$  (fail to reject)
- Conclusion: Insufficient evidence that mean differs from 500g



# From z-test to t-test

In most real-world situations:

- We **don't know** the population standard deviation  $\sigma$
- We only have sample data:  $\bar{x}$  and  $s$

## Question

What happens if we use  $s$  instead of  $\sigma$  in our test statistic?

$$\text{Incorrect: } z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

This doesn't follow  $N(0, 1)$ !

## From Normal to t-Student Distribution

When we replace  $\sigma$  with  $s$ , the distribution changes from normal to t-distribution

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

# t-Test Statistic for a Population Mean ( $\sigma$ Unknown)

## The t-Test Statistic Formula

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

where:

- $\bar{x}$ : Sample mean
- $\mu_0$ : Hypothesized population mean under  $H_0$
- $s$ : Sample standard deviation
- $n$ : Sample size
- $s/\sqrt{n}$ : Standard error of the sample mean  $\bar{x}$

## Sampling Distribution

Under  $H_0$ , the test statistic follows:

$$t \sim t_{n-1}$$

The t-distribution with  $n - 1$  degrees of freedom

# Comparison: z-test vs t-test

z-test ( $\sigma$ known)	t-test ( $\sigma$ unknown)
$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$
<b>Distribution:</b> $N(0, 1)$	<b>Distribution:</b> $t_{n-1}$
<b>Critical values:</b> From z-table	<b>Critical values:</b> From t-table
<b>Assumption:</b> $\sigma$ known	<b>Assumption:</b> $\sigma$ unknown
<b>Use when:</b> Rare in practice	<b>Use when:</b> Most common case

## Important

Always use  $t$ -test when  $\sigma$  is unknown, even for large samples!

- Some textbooks say "use  $z$  when  $n \geq 30$ "
- Modern practice: always use  $t$  when  $\sigma$  unknown
- $t$  converges to  $z$  as  $n$  increases anyway

# Critical Value Method: General Procedure

## 1 State hypotheses ( $H_0$ and $H_a$ )

Right tail  $H_0 : \mu \leq \mu_0$  vs  $H_a : \mu > \mu_0$

Left tail  $H_0 : \mu \geq \mu_0$  vs  $H_a : \mu < \mu_0$

Two tails  $H_0 : \mu = \mu_0$  vs  $H_a : \mu \neq \mu_0$

## 2 Check assumptions (random sample, normality, independence)

## 3 Compute test statistic

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

## 4 Determine critical value(s) based on:

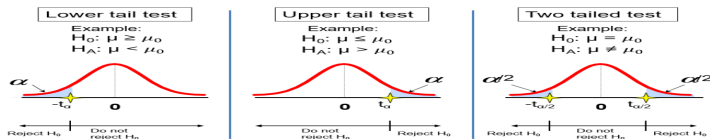
- ▶ Significance level  $\alpha$
- ▶ Type of test (one-tailed vs two-tailed)
- ▶ Degrees of freedom  $df = n - 1$

## 5 Make decision:

- ▶ Reject  $H_0$  if test statistic in rejection region
- ▶ Fail to reject  $H_0$  otherwise

## 6 State conclusion in context

# Critical Values for Different Test Types



Test Type	Rejection Region	Critical Value(s)	t-Table Lookup
Right-tailed test	$t > t_{\alpha, df}$	$t_{\alpha, df}$	Use column for $\alpha$
Left-tailed test	$t < -t_{\alpha, df}$	$-t_{\alpha, df}$	Use column for $\alpha$
Two-tailed test	$ t  > t_{\alpha/2, df}$	$\pm t_{\alpha/2, df}$	Use column for $\alpha/2$

df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$
5	1.476	2.015	2.571
10	1.372	1.812	2.228
15	1.341	1.753	2.131
20	1.325	1.725	2.086
25	1.316	1.708	2.060
30	1.310	1.697	2.042
$\infty$	1.282	1.645	1.960

Example:  $\alpha = 0.05$ ,  $df = 15$

- Right-tailed:  $t_{0.05, 15} = 1.753$
- Left-tailed:  $-t_{0.05, 15} = -1.753$
- Two-tailed:  $\pm t_{0.025, 15} = \pm 2.131$

## Symmetry Property

For one-tailed tests:

$$t_{\alpha, df} = -t_{1-\alpha, df}, \text{ where } P(T > t_{\alpha, df}) = \alpha$$

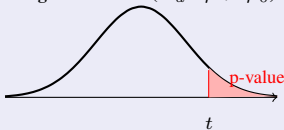
# p-Value Method: General Procedure

- 1 State hypotheses ( $H_0$  and  $H_a$ )
- 2 Compute test statistic  $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
- 3 Calculate p-value: Probability of obtaining results as extreme or more extreme than observed assuming  $H_0$  is true
- 4 Make decision:
  - ▶ Reject  $H_0$  if  $p\text{-value} \leq \alpha$
  - ▶ Fail to reject  $H_0$  if  $p\text{-value} > \alpha$
- 5 State conclusion in context

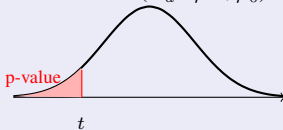
## Calculating p-Values for Different Test Types

Test Type	p-value Formula	Interpretation
$H_a : \mu > \mu_0$ (right-tailed)	$p\text{-value} = P(T \geq t)$	Area in right tail
$H_a : \mu < \mu_0$ (left-tailed)	$p\text{-value} = P(T \leq t)$	Area in left tail
$H_a : \mu \neq \mu_0$ (two-tailed)	$p\text{-value} = 2 \times P(T \geq  t )$	Area in both tails

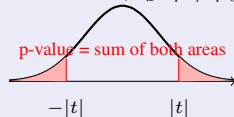
Right-tailed test ( $H_a : \mu > \mu_0$ )



Left-tailed test ( $H_a : \mu < \mu_0$ )



Two-tailed test ( $H_a : \mu \neq \mu_0$ )



# Example

## New Energy Drink Performance

A sports nutrition company introduces a new energy drink and claims that it leads to an increase in athletic performance relative to the current standard performance score of 75 points. To evaluate this claim at a significance level of  $\alpha = 0.05$ , a researcher takes a random sample of  $n = 25$  athletes who consumed the new energy drink. The sample mean performance score is  $\bar{x} = 82$  points, with a sample standard deviation of  $s = 15$  points. Based on this information, is there sufficient evidence to conclude that the new energy drink increases athletic performance compared to the standard?

### Solution:

#### Step 1: State the Hypotheses

The company claims an **increase** in performance, so we use a **right-tailed test**.

$$H_0 : \mu \leq 75$$

$$H_a : \mu > 75$$

where  $\mu$  is the population mean performance score.

## Step 2: Compute the Test Statistic

Since the population standard deviation  $\sigma$  is unknown, we use a **one-sample t-test**.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{82 - 75}{15/\sqrt{25}} = \frac{7}{3} \approx 2.33$$

Degrees of freedom:

$$df = n - 1 = 25 - 1 = 24$$

## Step 3: Critical Value Method

For a right-tailed test at  $\alpha = 0.05$  with  $df = 24$ , the critical value from the t-table is:

$$t_{\alpha, df} = t_{0.05, 24} = 1.711$$

Reject  $H_0$  if  $t > t_{\alpha, df}$

## Step 4: Decision (Critical Value Method)

Our Test Statistic:  $t = 2.333 > t_{\alpha, df} = 1.711$ , so we reject  $H_0$ .

## Step 5: Conclusion (Critical Value Method)

At the 5% significance level, there is sufficient evidence to conclude that the new energy drink increases athletic performance compared to the standard.



## Example: New Energy Drink Performance (Cont.)

### p-Value Method

For right-tailed test:

$$p\text{-value} = P(T \geq t) = P(T \geq 2.333), \quad \text{with } df = 24$$

Using t-Table at  $df = 24$ :

$t$ value	One-tailed $p$
$t_{0.025} = 2.064$	0.025
$t_{0.010} = 2.492$	0.010
2.333	between 0.01 and 0.025 which is less than $\alpha = 0.05$

Since  $0.01 < p\text{-value} < 0.025$  is less than  $\alpha = 0.05$ , we reject  $H_0$  and conclude that at the 5% significance level, there is sufficient evidence to conclude that the new energy drink increases athletic performance compared to the standard.

## Example: Coffee Temperature

A coffee shop claims their coffee is served at 75°C. A customer suspects it's different. They measure temperature of 12 cups:

73, 76, 74, 72, 77, 75, 74, 73, 76, 75, 74, 75

Test at  $\alpha = 0.05$  if the mean temperature differs from 75°C.

## Solution

### ① Hypotheses:

- ▶  $H_0 : \mu = 75$
- ▶  $H_a : \mu \neq 75$

### ② Sample Statistics: From the data (calculate or given):

- ▶ Sample size:  $n = 12$
- ▶ Sample mean:  $\bar{x} = \sum_{i=1}^n = 74.5$
- ▶ Sample standard deviation:  $s = \sqrt{\frac{\sum x^2 - (\sum x)^2/n}{n-1}} = 1.5$
- ▶ Degrees of freedom:  $df = n - 1 = 11$

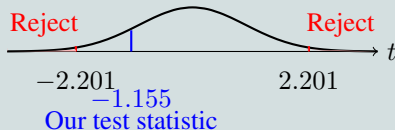
## Example: Coffee Temperature (Cont.)

### Test Statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{74.5 - 75}{1.5/\sqrt{12}} = \frac{-0.5}{1.5/3.464} = \frac{-0.5}{0.433} = -1.155$$

### Example: Critical Value Method

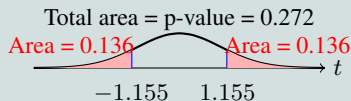
- 1 Significance level:  $\alpha = 0.05$
- 2 Critical value:  $t_{\alpha/2, df} = t_{0.025, 11} = 2.201$
- 3 Decision rule: Reject if  $|t| > 2.201$
- 4 Our test statistic:  $|t| = 1.155$
- 5 Since  $1.155 < 2.201$ , we fail to reject  $H_0$



## Example: Coffee Temperature (Cont.)

### p-value calculation

$p\text{-value} = 2 \times P(T \geq |1.155|)$ . From t-table with  $df = 11$ , we find  $t_{0.10,11} = 1.363$  ( $p = 0.2$ ) where our test statistic:  $t = -1.155$  is less than 1.363. Thus, we conclude that the test statistic has a corresponding p-value that is larger than  $p = 0.2$  and hence larger than  $\alpha = 0.05$ . So we fail to reject  $H_0$ .



### Conclusion for Both Methods

At  $\alpha = 0.05$ , we fail to reject  $H_0$ .

**Interpretation:** There is insufficient evidence to conclude that the mean coffee temperature differs from  $75^\circ\text{C}$ .

### Notes

- Both methods gave the same conclusion (as they always should!)
- We say "fail to reject  $H_0$ " not "accept  $H_0$ "
- The coffee might actually be different, but we don't have enough evidence with this sample

# Practice Problem: Textbook Weights (Try This Yourself)

A publisher claims their statistics textbook weighs 2.0 kg on average. You weigh 8 randomly selected textbooks:

1.9, 2.1, 2.0, 1.8, 2.2, 1.9, 2.0, 2.1

Test at  $\alpha = 0.05$  if the mean weight differs from 2.0 kg.

## Guided Steps

- 1 State hypotheses
- 2 Calculate  $\bar{x}$  and  $s$
- 3 Compute test statistic
- 4 Find critical value
- 5 Make decision
- 6 State conclusion

## Solution Check

- $\bar{x} = 2.0, s = 0.1414$
- $t = 0$  (exactly at null!)
- Critical value:  $t_{0.025,7} = 2.365$
- $|0| < 2.365$ : Fail to reject
- Conclusion: No evidence mean differs

# z Tests about a Population Proportion

We use a z-test for a population proportion when:

- We want to test a claim about a **population proportion** ( $p$ )
- Our data are **categorical** (success/failure, yes/no)
- We have a **single sample** from the population

## Examples:

- Testing if the proportion of voters supporting a candidate differs from 50%
- Testing if a drug's success rate is greater than 70%
- Testing if defect rate in a factory is less than 5%

# Conditions for z-Test for Proportion

Three conditions must be satisfied:

## Condition 1: Random Sample

The sample must be randomly selected from the population.

## Condition 2: Success-Failure Condition

Both  $np_0 \geq 5$  and  $nq_0 \geq 5$ , where:

- $n$  = sample size
- $p_0$  = null value (hypothesized population proportion) and  $q_0 = 1 - p_0$

## Condition 3: Independence

The sampled values must be independent of each other.

# The z-Test Statistic Formula

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Where:

- $\hat{p}$  = sample proportion =  $\frac{x}{n} = \frac{\text{number of successes in sample}}{\text{sample size}}$
- $p_0$  = hypothesized population proportion

**Note:**

- Denominator uses  $p_0$ , not  $\hat{p}$ !
- The denominator,  $\sqrt{\frac{p_0(1-p_0)}{n}}$ , is the *standard error* of  $\hat{p}$  under  $H_0$



# General Steps for Hypothesis Testing

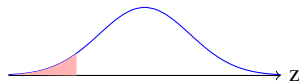
- ➊ **State** the null and alternative hypotheses
  - ▶ Right-tailed test:  $H_0 : p \leq p_0$  vs  $H_a : p > p_0$
  - ▶ Left-tailed test:  $H_0 : p \geq p_0$  vs  $H_a : p < p_0$
  - ▶ Two-tailed test:  $H_0 : p = p_0$  vs  $H_a : p \neq p_0$
- ➋ **Check** the conditions
- ➌ **Calculate** the test statistic
- ➍ **Make a decision** using either:
  - ▶ Critical value method, OR
  - ▶ P-value method
- ➎ **Interpret** the conclusion in context

# Critical Value Method: Step-by-Step

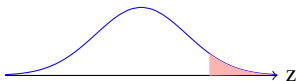
- ➊ Determine significance level  $\alpha$  (commonly 0.05)
- ➋ Identify the critical value(s) from z-table:
  - ▶ Right-tailed test:  $z_\alpha$
  - ▶ Left-tailed test:  $-z_\alpha$
  - ▶ Two-tailed test:  $\pm z_{\alpha/2}$
- ➌ Calculate test statistic  $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$ , where  $q_0 = 1 - p_0$
- ➍ Compare  $z$  to critical value(s)
- ➎ Decision rule:
  - ▶ If  $z > z_\alpha$  (right-tailed): Reject  $H_0$
  - ▶ If  $z < -z_\alpha$  (left-tailed): Reject  $H_0$
  - ▶ If  $|z| > z_{\alpha/2}$  (two-tailed): Reject  $H_0$

# P-Value Method: Step-by-Step

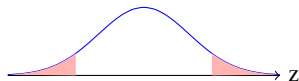
- 1 Calculate test statistic  $z$
- 2 Find the p-value:
  - ▶ Right-tailed test:  $p = P(Z > z)$
  - ▶ Left-tailed test:  $p = P(Z < z)$
  - ▶ Two-tailed test:  $p = 2 \times P(Z > |z|)$
- 3 Compare p-value to significance level  $\alpha$
- 4 Decision rule:
  - ▶ If  $p \leq \alpha$ : Reject  $H_0$
  - ▶ If  $p > \alpha$ : Fail to reject  $H_0$



Left-tailed:  $p = P(Z < z)$



Right-tailed:  $p = P(Z > z)$



Two-tailed:  $p = 2 \times P(Z > |z|)$

## Example: iPod Failure Rate

MacIn Touch reported that several versions of the iPod reported failure rates of 20% or more. From a customer survey, the colour iPod, first released in 2004, showed 64 failures out of 517. Is there any evidence that the failure rate for this model may be **lower than** the 20% rate of previous models? Assume  $\alpha = 0.001$ .

### Step 1: Null and Alternative Hypotheses

$$H_0 : p \geq 0.20 \quad \text{vs} \quad H_a : p < 0.20 \quad (\text{Left-tailed Test})$$

### Step 2: Check Success/Failure Condition

$$np_0 = 517 \times 0.20 = 103.4 > 5 \quad \checkmark$$

$$nq_0 = 517 \times 0.80 = 413.6 > 5 \quad \checkmark$$

$$\Rightarrow \hat{p} \sim N \left( p_0, \sqrt{\frac{p_0 q_0}{n}} \right)$$

### Step 3: Test Statistic Calculation

$$n = 517, \quad \hat{p} = \frac{x}{n} = \frac{64}{517} = 0.1238$$

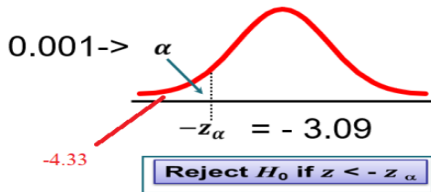
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.1238 - 0.20}{\sqrt{\frac{(0.20)(0.80)}{517}}} = -4.33$$

# iPod Failure Rate Example (Cont.)

## Step 4: Make a Decision (Critical Value Method)

### ► Using Critical Value Method:

- ★ Critical Value for  $\alpha = 0.001$ :  $z_{\alpha} = -3.09$
- ★ **Comparison:** Since  $z = -4.33 < z_{\alpha} = -3.09$ , we **reject the null hypothesis**.



### ► Using P-value Method:

$$P(Z < -4.33) \approx 0.0001$$

Since  $p\text{-value} < 0.0001 < \alpha = 0.001$ , we **reject the null hypothesis**.

## Conclusion

There is sufficient sample evidence to conclude that the failure rate for this iPod model is significantly lower than 20%.

## Example: Indigenous population proportion test

According to census data, Indigenous peoples make up approximately 3.3% of the total Canadian population. However, the proportion of Indigenous peoples is much higher in the territories. For example, in Nunavut, about 85% of the population is Indigenous.

To investigate this claim, a random sample of 300 residents in Nunavut is selected, and 261 individuals identify as Indigenous.

At a significance level of  $\alpha = 0.01$ , does the sample provide sufficient statistical evidence to conclude that the proportion of Indigenous peoples in Nunavut differs from the proportion reported in the national census data?

### Step 1: Null and Alternative Hypotheses

$$H_0 : p = 0.85 \quad \text{vs} \quad H_a : p \neq 0.85 \quad (\text{Two-tailed Test})$$

### Step 2: Check Success/Failure Condition

$$np_0 = 300 \times 0.85 = 255 > 5 \checkmark$$

$$nq_0 = 300 \times 0.15 = 45 > 5 \checkmark$$

$$\Rightarrow \hat{p} \sim N\left(p_0, \sqrt{\frac{p_0 q_0}{n}}\right)$$

### Step 3: Test Statistic Calculation

$$\begin{aligned} n &= 300, \hat{p} = \frac{x}{n} = \frac{261}{300} = 0.87 \\ z &= \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.87 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{300}}} = 0.97 \end{aligned}$$

# Indigenous proportion Example (Cont.)

**Step 4:** Using Critical Value Method: Reject  $H_0$  if  $z < -z_{\alpha/2} = -2.575$  or  $z > z_{\alpha/2} = 2.575$



Since  $-2.575 < z = 0.97 < 2.575$ , we fail to reject  $H_0$ .

**Step 4:** Using P-value Method:

$$\begin{aligned} 2P(Z > 0.97) &= P(Z < -0.97) + P(Z > 0.97) \\ &= 2(1 - 0.8340) = 0.332 \end{aligned}$$

Since  $p\text{-value} = 0.332 > \alpha = 0.01$ , we fail to reject  $H_0$ .

**Step 5:** Conclusion: We conclude that the proportion of indigenous people in Nunavut is NOT different from that reported in census data.

# Practice Problem: Textbook Weights (Try This Yourself)

A pharmaceutical company claims their vaccine is 90% effective. In a clinical trial with 200 patients, 168 were successfully immunized. Is there evidence that the effectiveness is different from 90%? Use  $\alpha = 0.05$ .

## Guided Steps

- 1 State hypotheses
- 2 Check the assumptions
- 3 Calculate  $\hat{p}$
- 4 Compute test statistic
- 5 Find critical value
- 6 Make decision
- 7 State conclusion

## Solution Check

- $H_0 : p = 0.90$  vs.  $H_a : p \neq 0.90$
- $\hat{p} = \frac{168}{200} = 0.84$
- $z \approx -2.83$
- Critical values are  $\pm 1.96$
- Since  $|z| = 2.83 > 1.96$ , we reject  $H_0$
- The p-value is approximately  $0.0046 < 0.05$ , so we reject  $H_0$ .
- At the 5% significance level, there is sufficient evidence to conclude that the vaccine effectiveness is different from 90%.



# Summary of One-Sample Test Statistics

## Population Mean $\mu$

**z-test ( $\sigma$  known):**

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

**Conditions:**

- $\sigma$  is known
- $n \geq 30$  OR population normal

**t-test ( $\sigma$  unknown):**

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad df = n - 1$$

**Conditions:**

- $\sigma$  is unknown

## Population Proportion $p$

**z-test:**

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

where  $\hat{p} = \frac{x}{n}$  **Conditions:**

- 1 Random sample
- 2 Independent sample
- 3  $np_0 \geq 5$
- 4  $nq_0 \geq 5$ , where  $q_0 = 1 - p_0$

## Two Decision Methods:

- **Critical value:** Compare test-statistic to critical value (tabular value) based on  $\alpha$  and  $H_0$
- **P-value:** Compare p-value to  $\alpha$  and reject  $H_0$  if p-value  $< \alpha$ .