PART A:
Hypothesis Testing

LEARNING OBJECTIVES

At the end of this section, you should be able to do the following:

- Formulate **null** and **alternative hypotheses** for applications involving a **single population** mean or proportion.
- Know what **Type I** and **Type II errors** are.
- Correctly formulate a **decision rule** for testing a hypothesis.
- Know how to use the **test statistic**, **critical value**, and **p-value** approaches to test a hypothesis.
- **Conduct Hypothesis testing** involving a single population parameter in Excel (using Hypothesis testing templates).

THE TWO KEY INFERENTIAL TOOLS

Confidence interval

- Use confidence interval estimation if we have no idea about the value of the population parameter being investigated.

Hypothesis test

- Use hypothesis tests when we have some idea of the value of the population parameter being investigated, or if we have some hypothesised value against which we can compare our sample results.

CASE STUDY
SOCIAL MEDIA MARKETING

- What number of ads do we need to display on one's Facebook page before s/he starts following our page?

OR

- The mean number of Facebook ads required to turn a fan into a follower is more than five.

SIX STEPS IN HYPOTHESIS TESTING

1	Set up H_0 and H_1	Describe in words and symbols
2	Decide on the direction of the test	Two tail, Lower tail, Upper tail
3	Choose α, and sampling H_0 true, determine the critical value(s) of the test statistic	How serious is a Type I error?
4	Decision Rule	Use critical values of test statistic
5	Observed (sample) value of test statistic assuming H_0 true	Perform relevant calculations
6	Conclusion	In terms of H_0 and H_1 and in practical terms

1. FORMULATING THE HYPOTHESES

Null Hypothesis

- The statement about the population parameter that will be assumed to be true during the conduct of the hypothesis test.
- The null hypothesis will be rejected only if the sample data provide substantial contradictory evidence.

Alternative Hypothesis

- The hypothesis that includes all population values not included in the null hypothesis.

FORMULATING HYPOTHESIS

EXAMPLES

Testing the ***Status Quo***

- The box of cereal has a mean fill of 16 ounces.

$$H_0: \mu = 16$$
$$H_A: \mu \neq 16$$

Testing a Research ***Hypothesis***

- Goodyear's tyre will last longer than its competitor's on average; more than 60,000 km.

$$H_0: \mu \leq 60{,}000$$
$$H_A: \mu > 60{,}000$$

Testing a ***Claim*** about Population

- Average waiting time in a medical clinic is less than 15 minutes.

$$H_0: \mu \geq 15$$
$$H_A: \mu < 15$$

WHAT ARE THE HYPOTHESES IN THE CASE STUDY?

- The mean number of Facebook ads required to turn a fan into a follower is more than five.

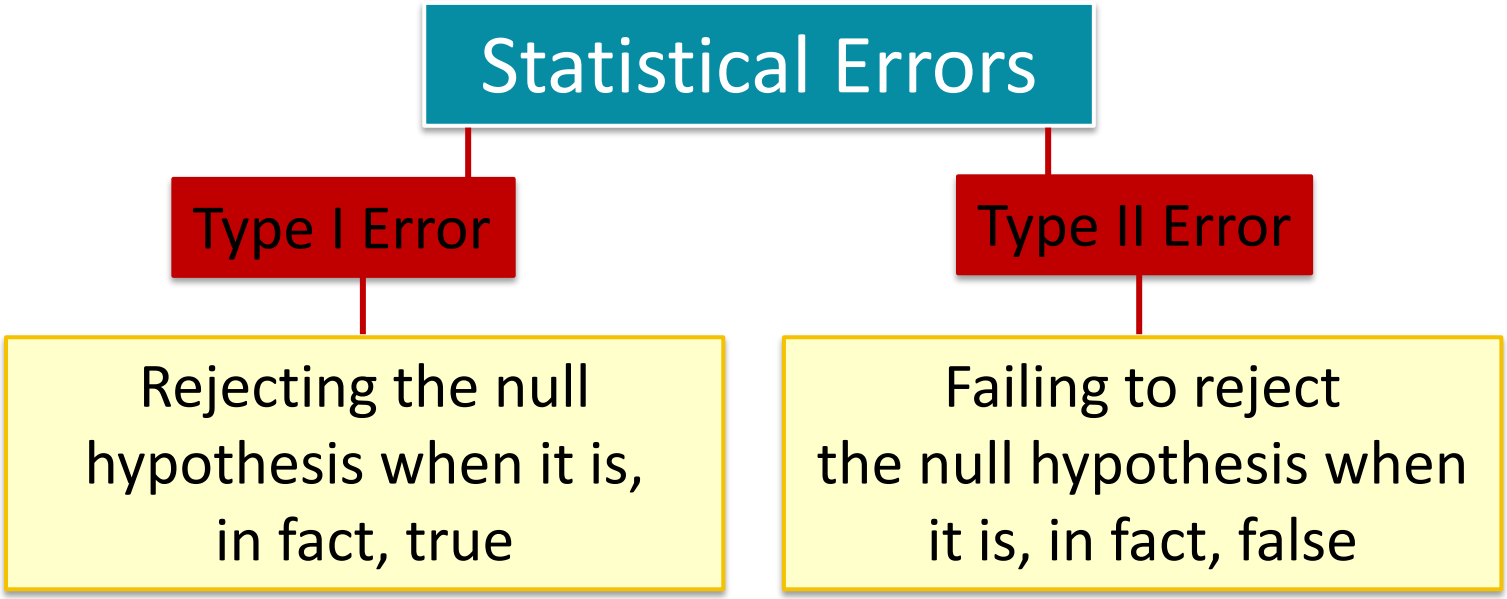
H_0: The mean number of Facebook ads required to turn a fan into a follower is less than or equal to five.

$H_0: \mu \leq 5$

H_1: The mean number of Facebook ads required to turn a fan into a follower is greater than five.

$H_1: \mu > 5$

TYPES OF STATISTICAL ERRORS



		State of Nature	
		Null Hypothesis True	Null Hypothesis False
Decision	Conclude Null True (Don't Reject H_0)	Correct Decision	Type II Error
	Conclude Null False (Reject H_0)	Type I Error	Correct Decision

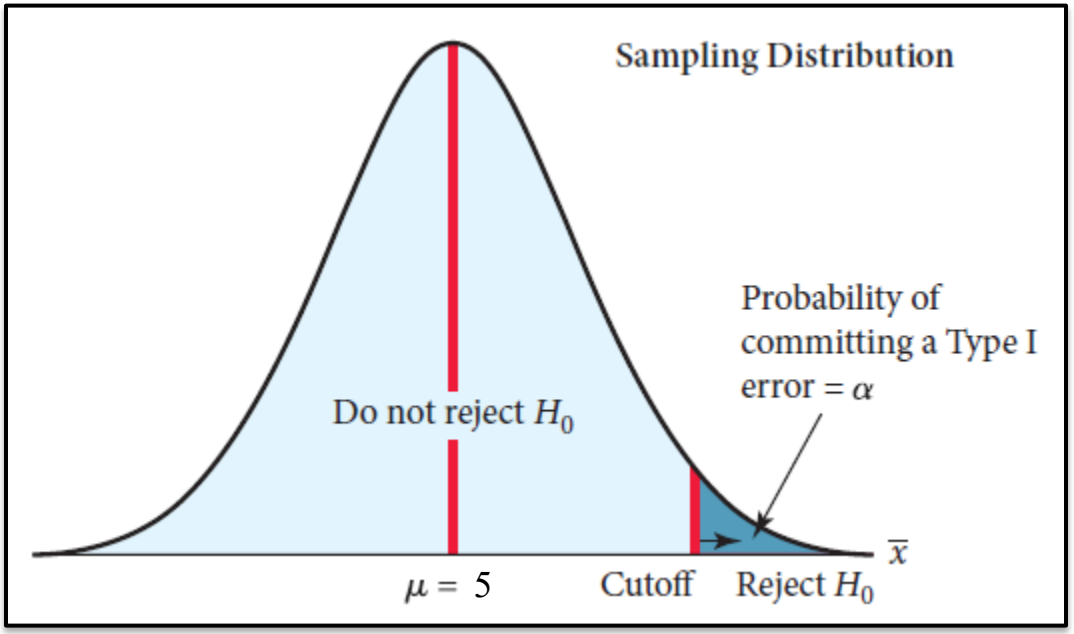
TYPE I AND II ERRORS

Type I Error

Type II Error

SIGNIFICANCE LEVEL AND CRITICAL VALUE

$H_0: \mu \leq 5$
$H_1: \mu > 5$



- Significance Level (α)

The maximum allowable probability of committing Type I error.

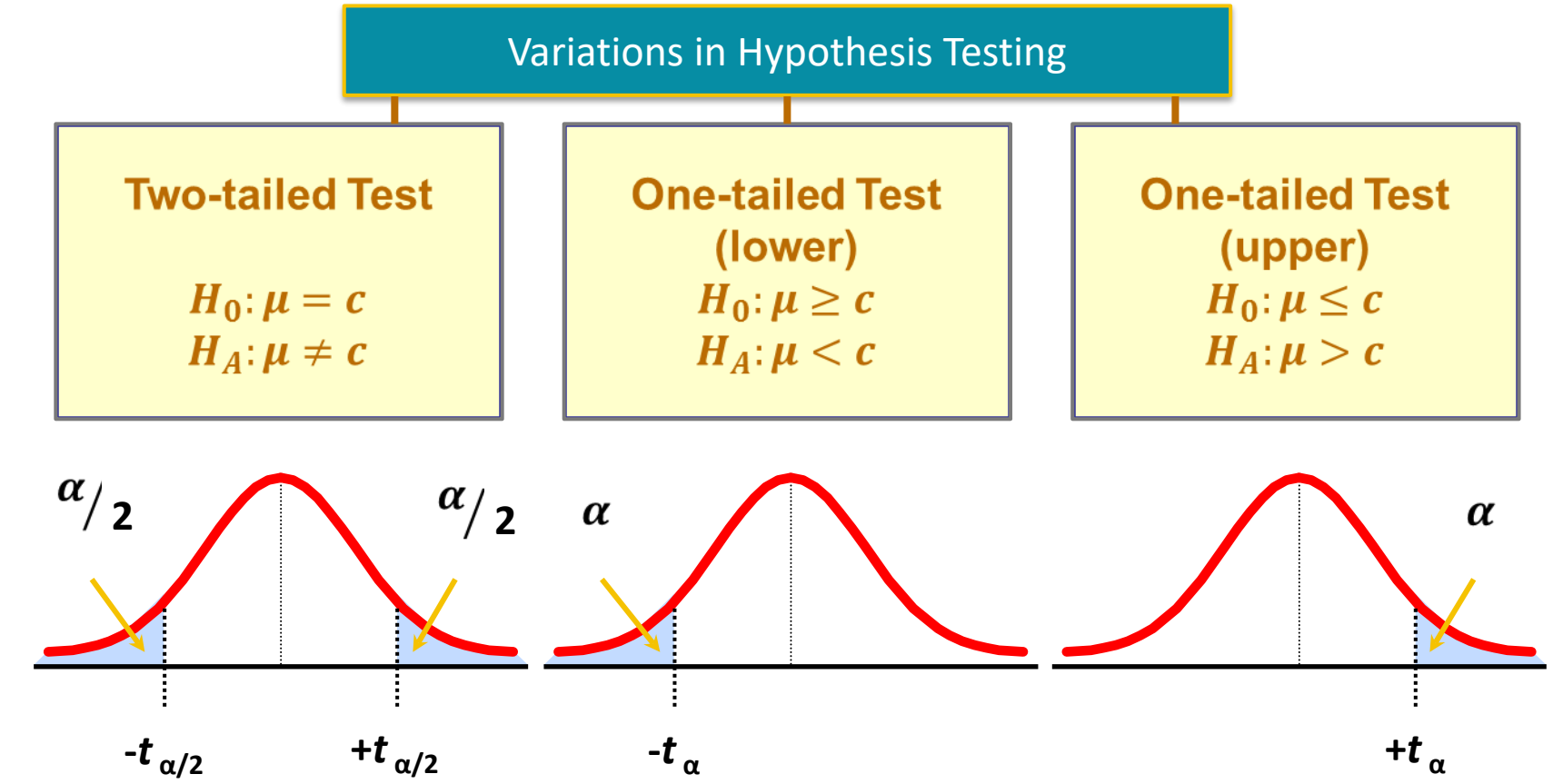
SIGNIFICANCE LEVEL AND CRITICAL VALUE

- The value of α is determined based on the costs involving in committing type I error.

Critical value:

- The value corresponding to a significance level that determines those test statistics that lead to rejecting the null hypothesis and those that lead to a decision not to reject.

2. TYPES OF H-TEST



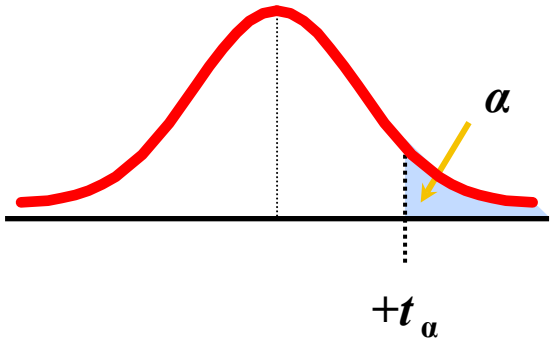
WHAT TYPE OF TEST REQUIRED IN THE CASE STUDY?

- The mean number of Facebook ads required to turn a fan into a follower is more than five.

$H_0: \mu \leq 5$

$H_1: \mu > 5$

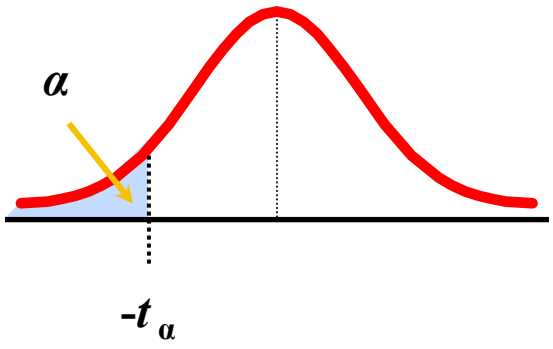
- One tail (Upper-tail) test



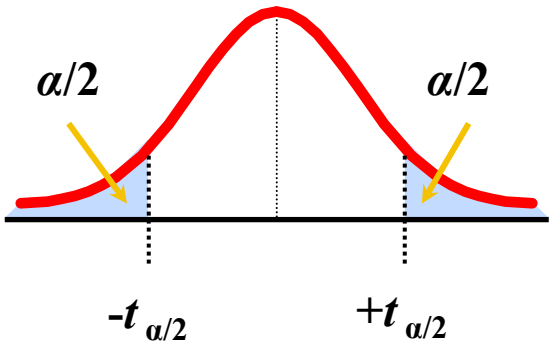
4. DECISION RULE

- If our sample value of the test statistic falls beyond the critical value(s), we reject H_0 in favour of H_1

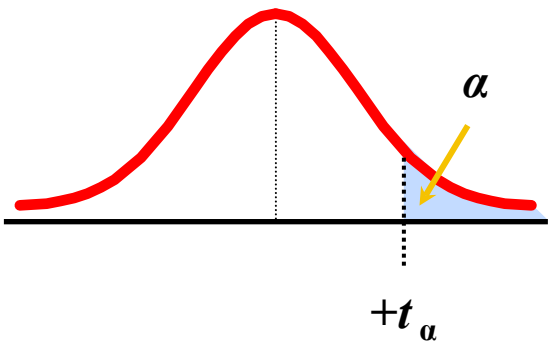
If the sample value of the test statistic < critical value we reject H_0 in favor of H_1



If the sample value of the test statistic lies in either critical region, we reject H_0 in favor of H_1



If the sample value of the test statistic > critical value we reject H_0 in favor of H_1



- For the case study, we reject H_0 if the sample t-statistic > +1.645

5. SAMPLE ANALYSIS (CALCULATING TEST STATISTIC)

- Collect random sample
- Calculate sample statistics
- Use sample mean for hypothesis test of mean
- Use sample proportion *p* for hypothesis tests of proportion
- Calculate the sample value of the test statistic, *t-statistic*

- *t-Test Statistic:*

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

Where:

$\bar{x}$ = sample mean
μ = hypothesised value of the population mean
s = sample standard deviation
n = sample size

TEST STATISTIC FOR THE CASE STUDY

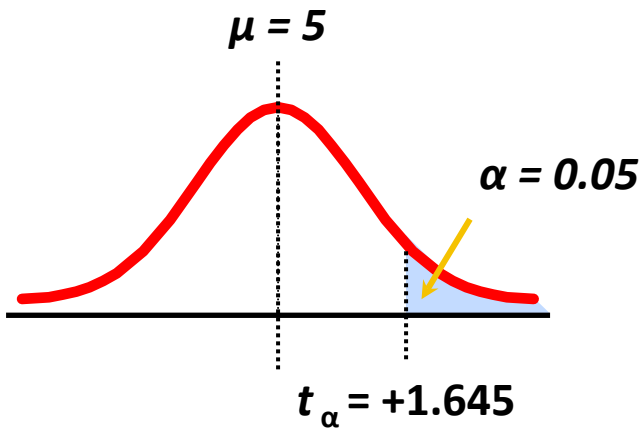
- Let's assume

$\bar{x}$ = 5.2

μ = 5

s = 1.2

n = 300

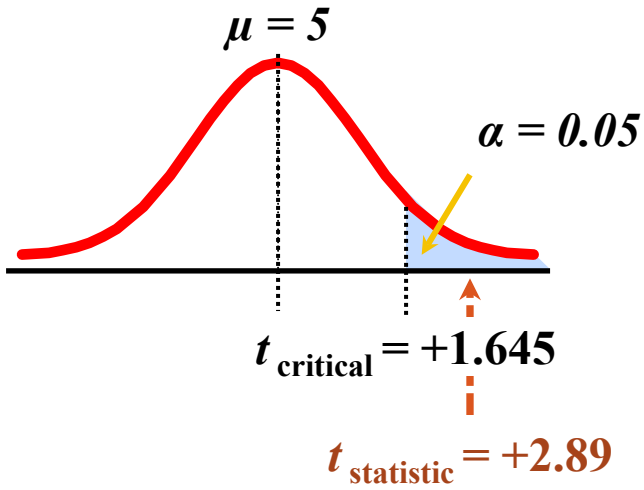


$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \qquad t = \frac{5.2 - 5}{\frac{1.2}{\sqrt{300}}} = +2.89$$

6. DRAW CONCLUSION

- Compare the sample value of the test statistic (from step 5) to your decision rule (from step 4)

  For the case study, the sample statistic of t = +2.89 > critical value of +1.645 and therefore lies in the rejection region. We reject H_0.



  At 5 percent significance, there is sufficient evidence to conclude that the true average number of Facebook ads required to turn a fan into a follower is greater than five.

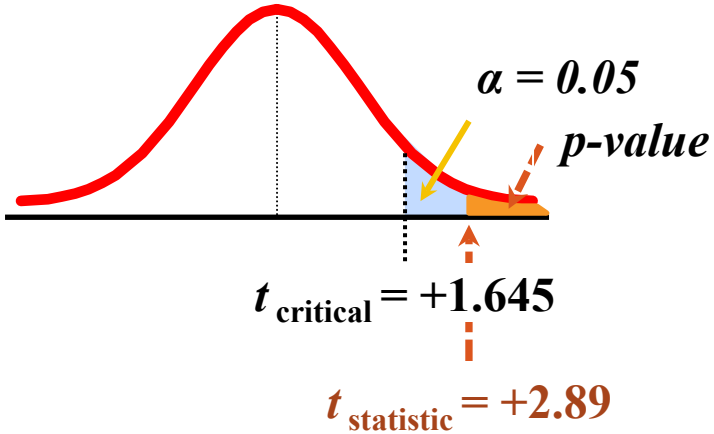
P-VALUE APPROACH TO H-TEST

p-Value

- The probability (assuming the null hypothesis is true) of obtaining a test statistic at least as extreme as the test statistic we calculated from the sample. The p-value is also known as the observed significance level.

If p-value < α, reject H_0

If p-value ≥ α, do not reject H_0



WHY P-VALUE APPROACH

- It's more straightforward to draw a conclusion using the p-value approach.
- Adds a degree of significance to the result of the hypothesis test
- More than just a simple "reject"
  - Can now determine how strongly you "reject" or "accept"
- The farther the p-value is from α, the stronger the decision
- Almost all analytics tools calculate p-value.

CASE STUDY
SOCIAL MEDIA MARKETING

- What proportion HBO Facebook followers have subscribed to HBO GO (online video streaming service)?

OR

- Less than 15% of all HBO Facebook followers have subscribed to its online video streaming service.

HYPOTHESIS TESTING FOR PROPORTIONS

- Involves categorical values
- Two possible outcomes
- More than just a simple “reject”
  - “Success” (possesses a certain characteristic)
  - “Failure” (does not possesses that characteristic)

- Example:
  - Defective items manufactured

HYPOTHESIS TESTING FOR PROPORTIONS

1. Specify the population parameter of interest (i.e. π) and the null and alternative hypotheses
2. Type of test
3. Specify the significance level and determine the critical z value.
4. Decision rule

$$p = \frac{x}{n}$$

5. Compute the test statistic (or the p-value)

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

6. Reach decision by comparing sample $z_{statistic}$ to $z_{critical}$ or p-value to α, draw conclusion and interpret the findings

WHAT ARE THE HYPOTHESES IN THE CASE STUDY?

- Less than 15% of all HBO Facebook followers have subscribed to its online video streaming service.

H_0: The true proportion of all HBO Facebook followers who have subscribed to its online video streaming service is greater than or equal to 15%.

$H_0: \pi \geq 15\%$

H_1: The true proportion of all HBO Facebook followers who have subscribed to its online video streaming service is less than 15%.

$H_1: \pi < 15\%$

H-TESTING (PROPORTION)

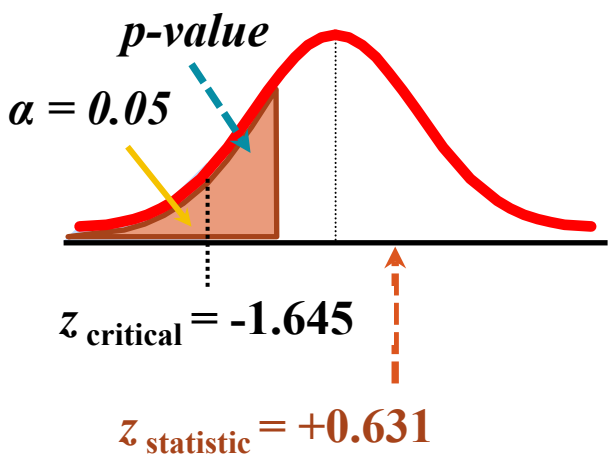
- Type of Test - Lower tail test
- Assume 5% significance, sample size = 300, number of subscribers in the sample = 49.
- $Z_{critical} = -1.645$
- If $z_{statistic} < z_{critical}$, then reject H_0. Otherwise do not reject H_0.
- $Z_{statistic} = \frac{0.163-0.15}{\sqrt{\frac{0.15(1-0.15)}{300}}} = +0.631$

H-TESTING (PROPORTION)

- For the case study, the sample statistic of z = +0.631 > critical value of -1.645 and therefore lies in the non-rejection region. We do NOT reject H_0.

  At 5 percent significance level, there is no sufficient evidence to conclude that less than 15% of all HBO Facebook followers have also subscribed to its online video streaming service.

p-value : 0.7357



QUESTIONS?