HYPOTHESIS TESTING

INFERENTIAL STATISTICS

prepared by:

Gyro A. Madrona

Electronics Engineer

TOPIC OUTLINE

Hypothesis Test

Rejection Region

Critical Value and Z-score

p-Value

HYPOTHESIS TEST

HYPOTHESIS

A **hypothesis** is an initial **assumption** formed before collecting data, and it serves as a statement about a **population** parameter rather than about the sample data.

HYPOTHESIS TEST

A **hypothesis test** is simply comparing reality to an assumption and asking, "**Did things change?**"

Null Hypothesis (H_o)

Represents **no change**, no effect, or the status quo.

Alternative Hypothesis (H_a)

Represents the possibility that things did change or that there is a **significant difference**.

IS YOUR DATA GUILTY?

Hypothesis testing is like a legal system where the defendant is assumed **innocent** until proven guilty.

The Truth

H_o: Innocent

	Innocent	Guilty
Accept	✓	✗
Reject	✗	✓

TYPES OF ERROR

1. **Type I Error**

Rejecting the null hypothesis when it is actually true. The risk of making type I error is denoted by $\boldsymbol{\alpha}$ (e.g., 0.05).

2. **Type II Error**

Failing to reject the null hypothesis when it is actually false. The risk of making a type II error is denoted by $\boldsymbol{\beta}$ (e.g. 0.20)

The Truth

H_o: Innocent	Innocent	Guilty
Accept	✓	**Type II Error** (False negative)
Reject	**Type I Error** (False positive)	✓

EXERCISE

A company claims that the average lifespan of their batteries is 500 hours. A consumer group suspects that the average lifespan is **different** from **500 hours**.

Null Hypothesis

$$H_o: \mu_1 = 500$$

The average battery lifespan is 500 hours

Alternative Hypothesis

$$H_a: \mu_1 \neq 500$$

The average battery lifespan differs from 500 hours

EXERCISE

A company claims that the average lifespan of their batteries is 500 hours. A consumer group suspects that the batteries last **fewer than 500 hours**.

Null Hypothesis

$$H_o\colon \mu_1 \geq 500$$

The average battery lifespan is at least 500 hours

Alternative Hypothesis

$$H_a\colon \mu_1 < 500$$

The average battery lifespan is fewer than 500 hours

EXERCISE

A company claims that the average lifespan of their batteries is 500 hours. An independent lab believes that the batteries last **longer than 500 hours**.

Null Hypothesis

$$H_o: \mu_1 \leq 500$$

The average battery lifespan is 500 hours at most

Alternative Hypothesis

$$H_a: \mu_1 > 500$$

The average battery lifespan is longer than 500 hours

REJECTION REGION

SIGNIFICANCE LEVEL

The **significance level** ($\boldsymbol{\alpha}$) determines the threshold for deciding whether to **reject** the null hypothesis ($\boldsymbol{H_o}$).

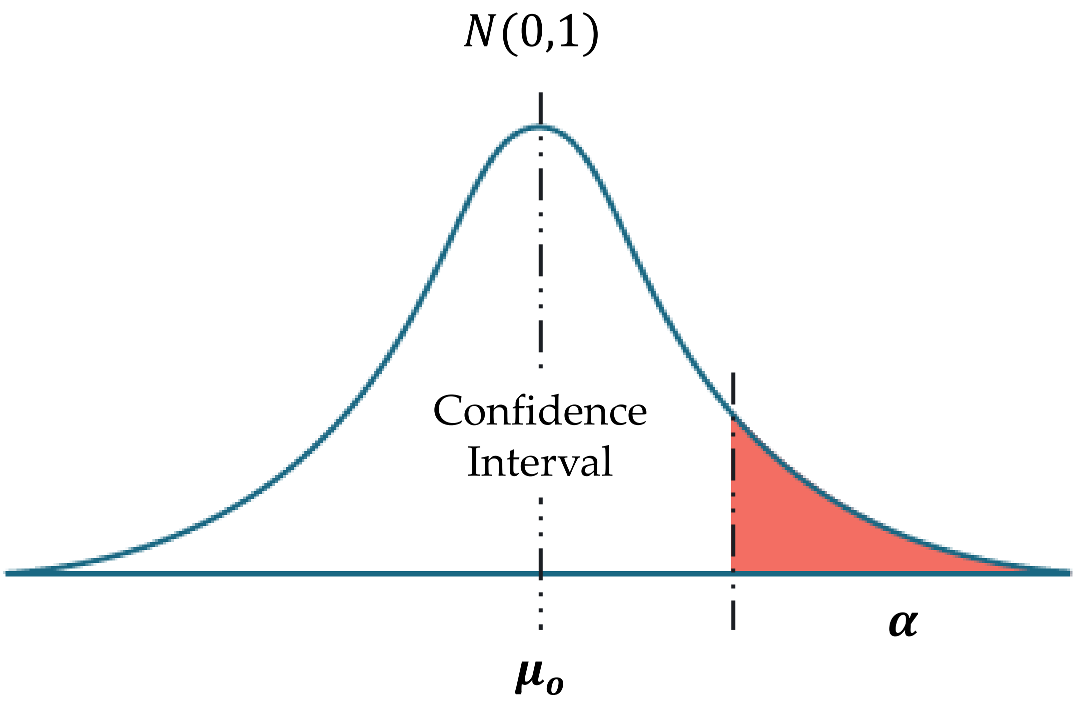
Typical values for $\boldsymbol{\alpha}$

0.01

0.05

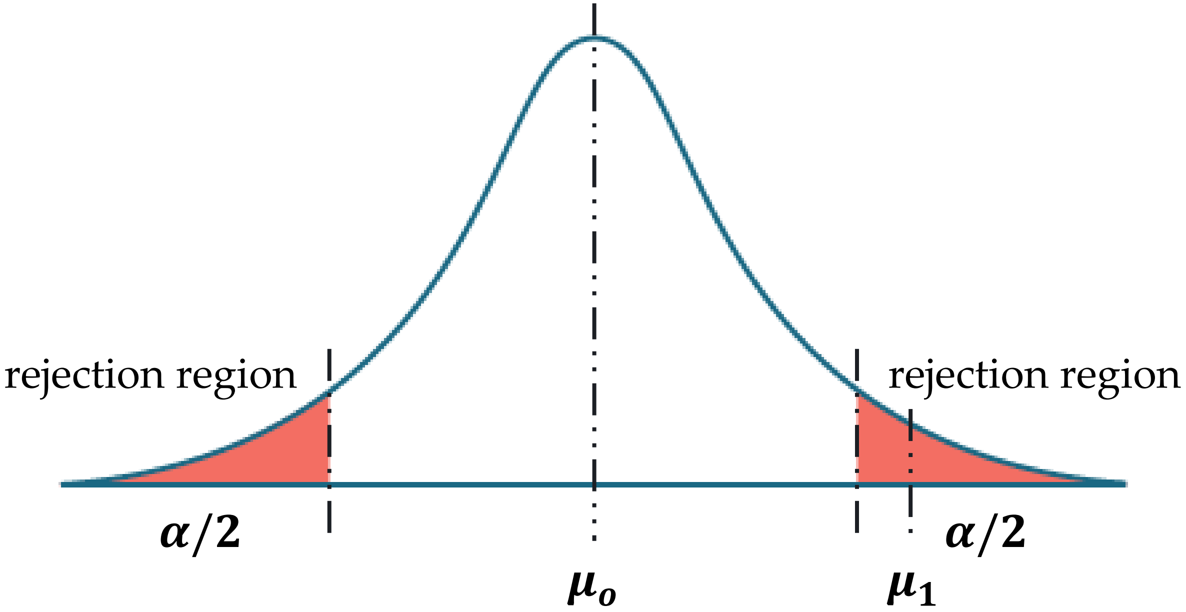
0.1

Standard Normal Distribution



REJECTION REGION

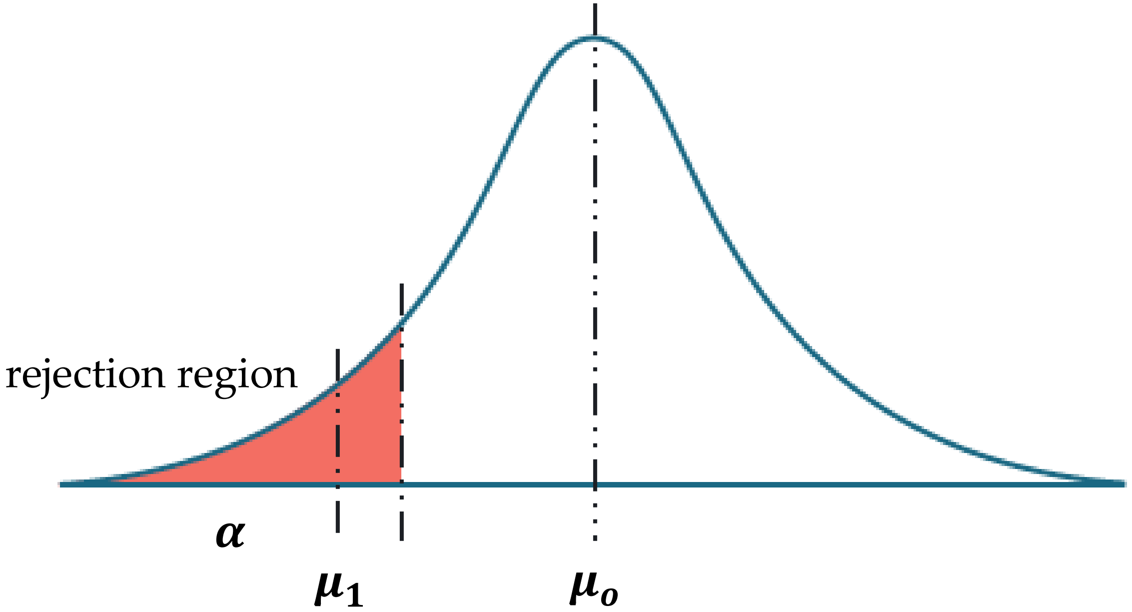
Two-Tailed Test



H_o: $\mu_1 = \mu_o$

H_a: $\mu_1 \neq \mu_o$

Left-Tailed Test

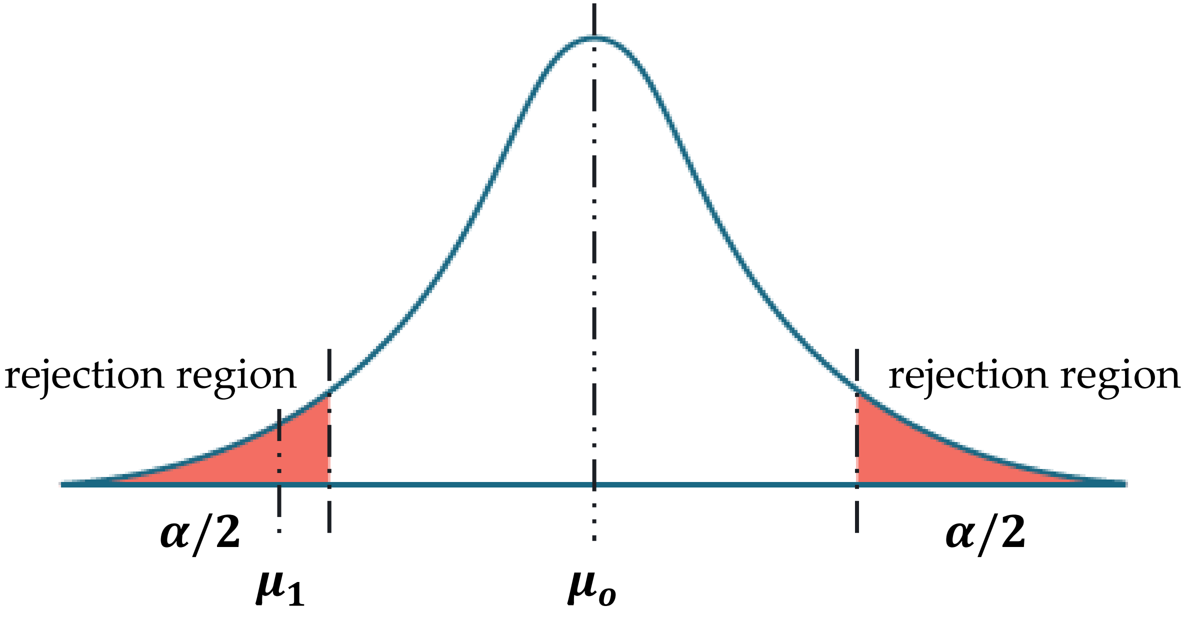


H_o: $\mu_1 = \mu_o$

H_a: $\mu_1 < \mu_o$

REJECTION REGION

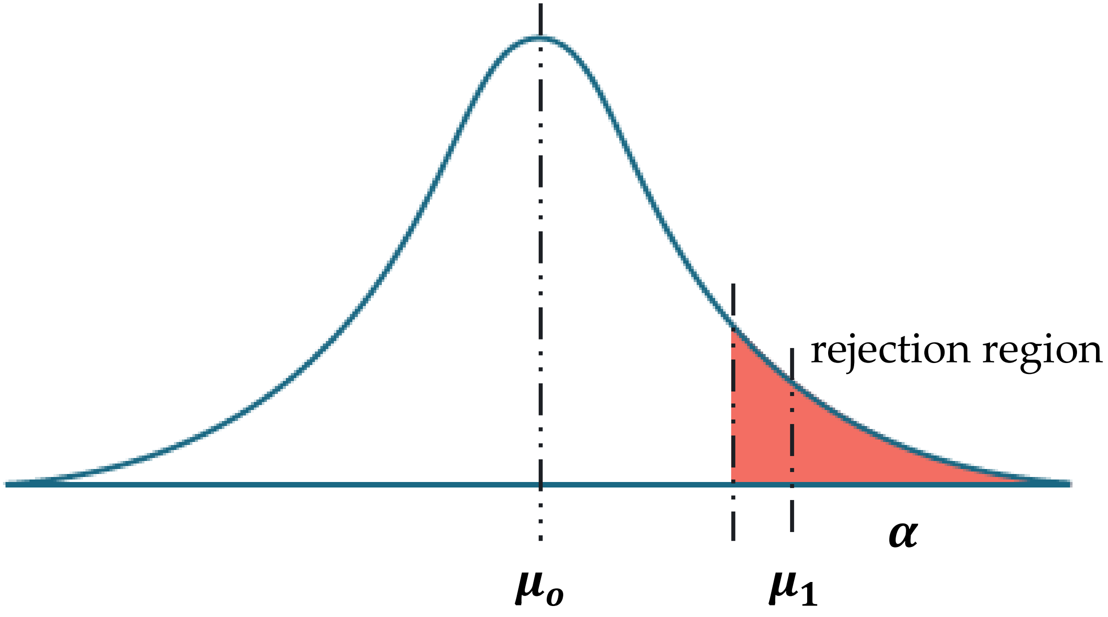
Two-Tailed Test



H_o: $\mu_1 = \mu_o$

H_a: $\mu_1 \neq \mu_o$

Right-Tailed Test



H_o: $\mu_1 = \mu_o$

H_a: $\mu_1 > \mu_o$

CRITICAL VALUE AND Z-SCORE

CRITICAL VALUE AND Z-SCORE

lowercase z

z refers to the **critical value** obtained from the standard normal distribution table (z-table).

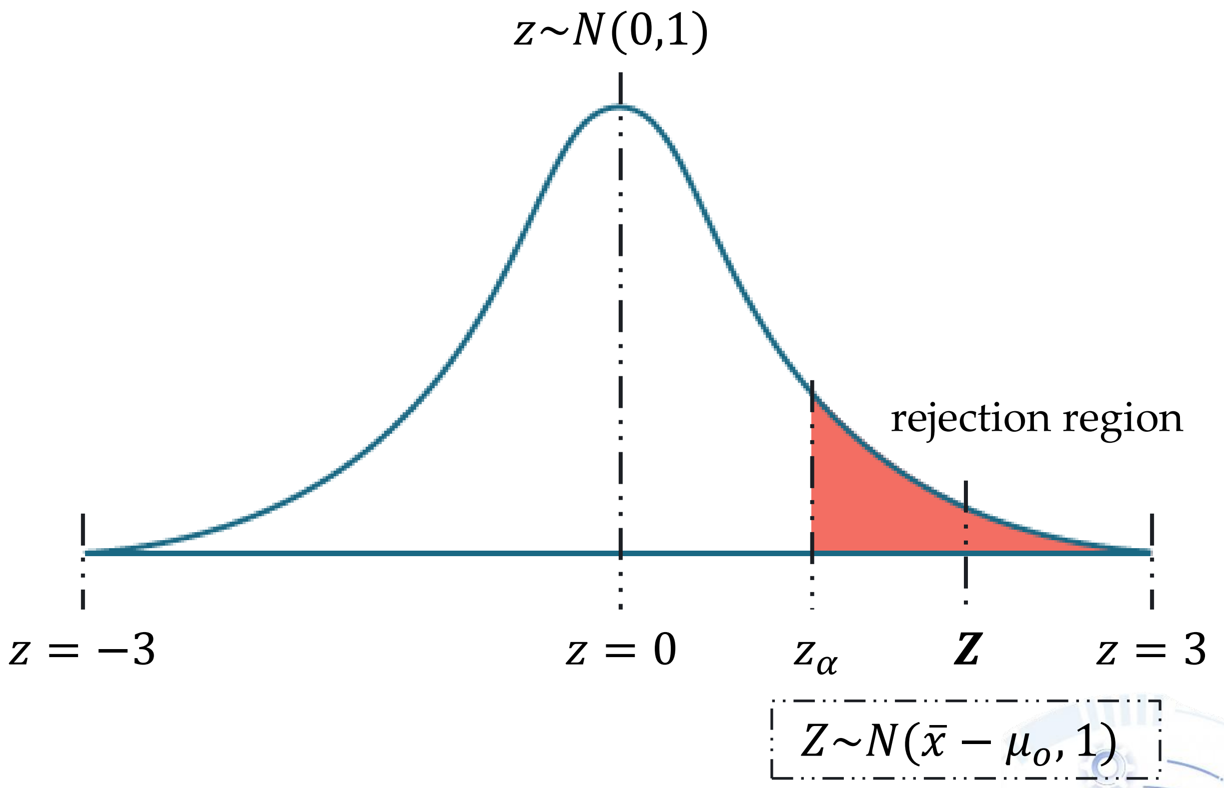
uppercase Z

Z is a standardized variable associated with the test called the **Z-score**.

Formula

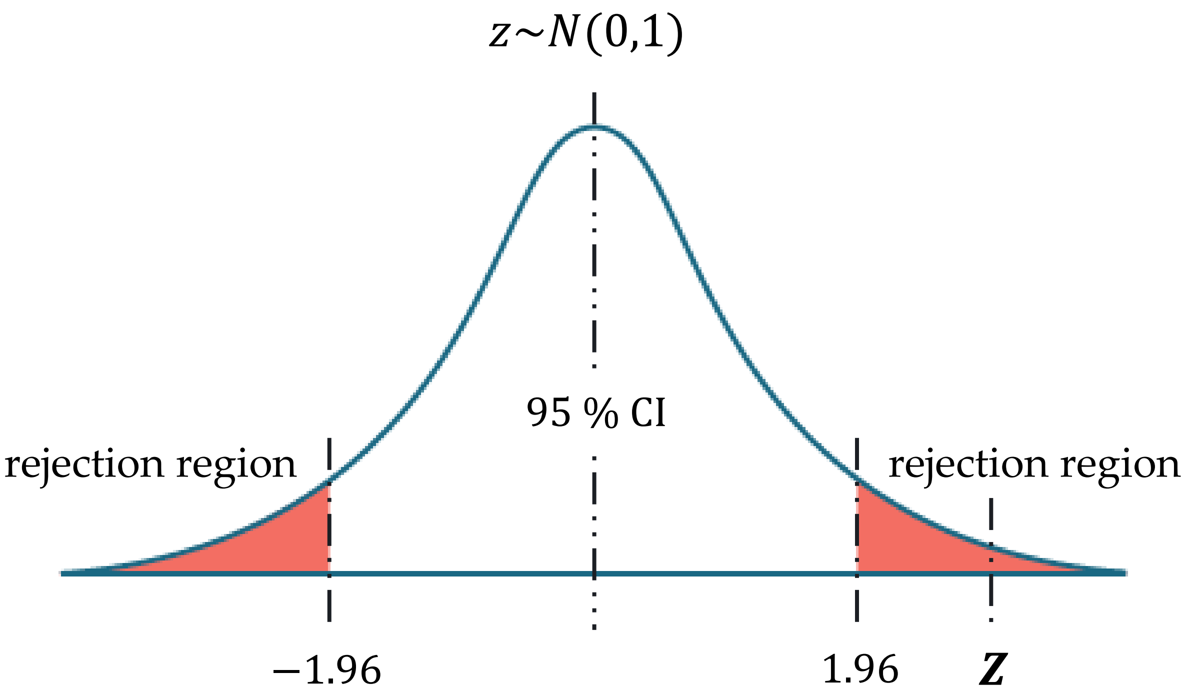
$$Z = \frac{\bar{x} - \mu_o}{\sigma/\sqrt{n}}$$

Right-Tailed Test



EXERCISE

Two-Tailed Test



$\alpha = 0.05$

$z_{0.025} = 1.96$

Null Hypothesis

$\boldsymbol{H_o: \mu_1 = 500}$

The average battery lifespan is 500 hours

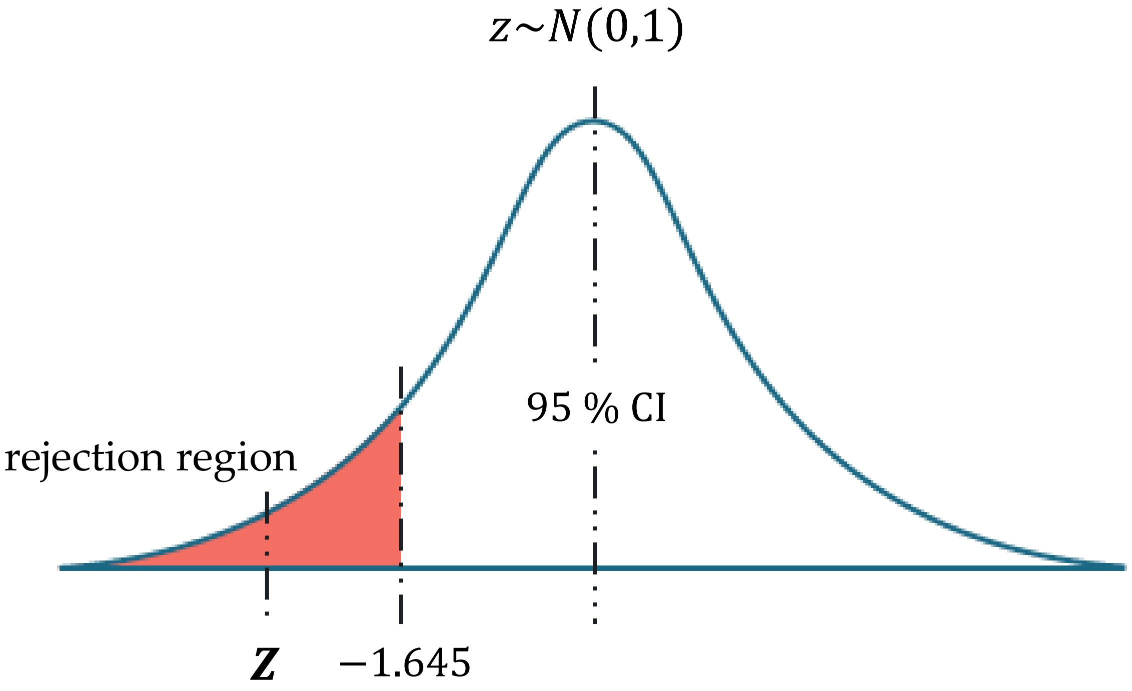
Alternative Hypothesis

$\boldsymbol{H_a: \mu_1 \neq 500}$

The average battery lifespan differs from 500 hours

EXERCISE

Left-Tailed Test



$\alpha = 0.05$

$z_{0.05} = 1.645$

Null Hypothesis

H_o: $\mu_1 \geq 500$

The average battery lifespan is at least 500 hours

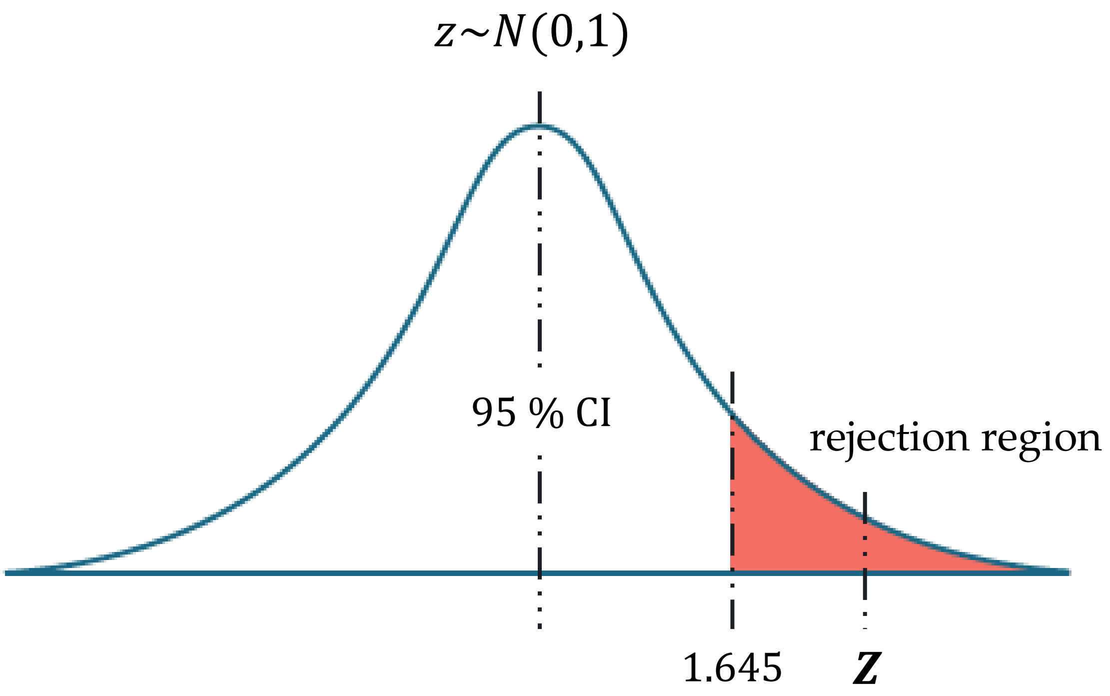
Alternative Hypothesis

H_a: $\mu_1 < 500$

The average battery lifespan is fewer than 500 hours

EXERCISE

Right-Tailed Test



$\alpha = 0.05$

$z_{0.05} = 1.645$

Null Hypothesis

$\boldsymbol{H_o}$: $\boldsymbol{\mu_1 \leq 500}$

The average battery lifespan is 500 hours at most

Alternative Hypothesis

$\boldsymbol{H_a}$: $\boldsymbol{\mu_1 > 500}$

The average battery lifespan is longer than 500 hours

EXERCISE

A manufacturing process is claimed to have an average defect rate of **10.32** units, with a known standard deviation of **3.17** units. The Statistical Process Control (SPC) department suspects this claim may no longer be valid and collects a **random sample** of **30** production units to test whether the true average **defect rate differs** significantly from **10.32**.

dataset

"**defects-30-sample.csv**"

Solution

EXERCISE

A manufacturing process is claimed to have an average defect rate of **10.32** units, with a known standard deviation of **3.17** units. The Statistical Process Control (SPC) department suspects this claim may no longer be valid and collects a **random sample** of **30** production units to test whether the true average **defect rate increases** significantly from **10.32**.

dataset

"**defects-30-sample.csv**"

Solution

P-VALUE

P-VALUE

The **p-value** (probability value) is the **smallest level of significance** at which we can still reject the null hypothesis, given the observed sample statistic.

One-Tailed Test

$$\textbf{p-value} = \mathbf{1} - \text{value from the table}$$

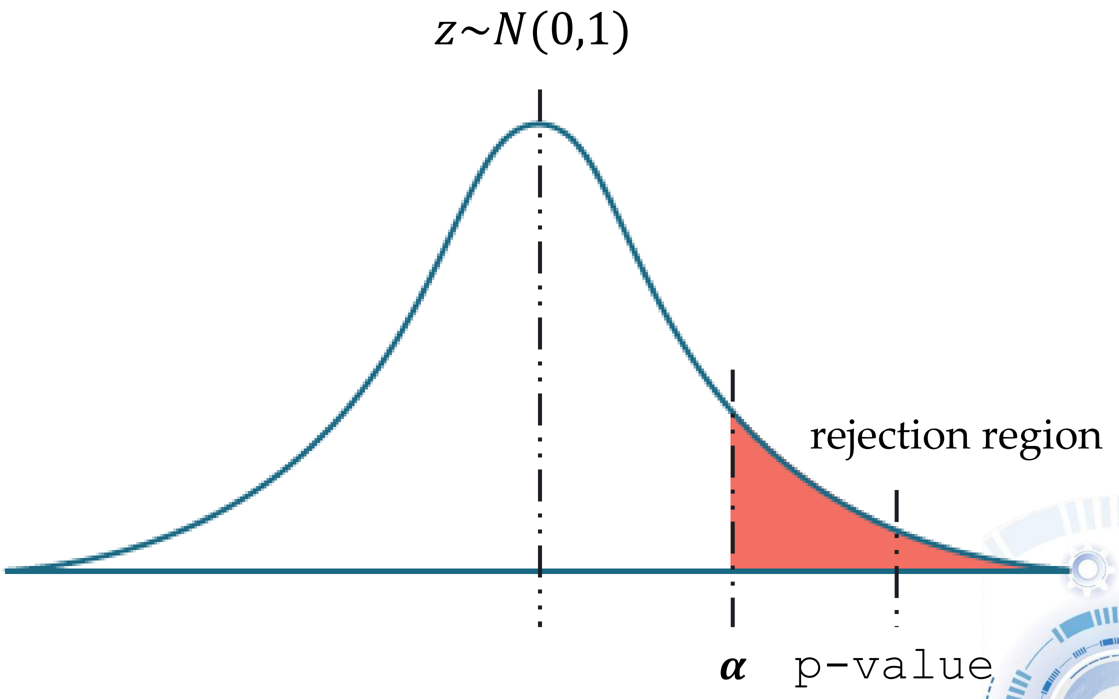
Two-Tailed Test

$$\textbf{p-value} = (\mathbf{1} - \text{value from the table})\ \times\mathbf{2}$$

Hypothesis Test

Reject H_o if **p-value** $< \boldsymbol{\alpha}$

Fail to reject H_o if p-value $\geq \alpha$



P-VALUE

The **p-value** (probability value) is the **smallest level of significance** at which we can still reject the null hypothesis, given the observed sample statistic.

One-Tailed Test

p-value = 1 − value from the table

Two-Tailed Test

p-value = (**1** − value from the table) ×**2**

syntax

```
from scipy import stats
```

One-Tailed Test

```
p_value = 1-stats.norm.cdf(Z_score)
```

Two-Tailed Test

```
p_value = 2*(1-stats.norm.cdf(Z_score))
```

EXERCISE

A manufacturing process is claimed to have an average defect rate of **10.32** units, with a known standard deviation of **3.17** units. The Statistical Process Control (SPC) department suspects this claim may no longer be valid and collects a **random sample** of **30** production units to test whether the true average **defect rate differs** significantly from **10.32**.

dataset

"**defects-30-sample.csv**"

Solution

LABORATORY