Moringa School |

www.moringaschool.com

Hypothesis Testing

Random callouts

Content Walkthrough

Hypothesis Testing

Objective

- Identify the steps of hypothesis testing.
- Define the null hypothesis, alternative hypothesis, level of significance, test statistic, p-values, and statistical significance.
-

Introduction

Hypothesis testing is a statistical method used to make decisions about the population based on sample data. It involves making an assumption (the **null hypothesis**, H_0) and determining whether the data provide enough evidence to reject this assumption in favor of an alternative hypothesis (H_1)

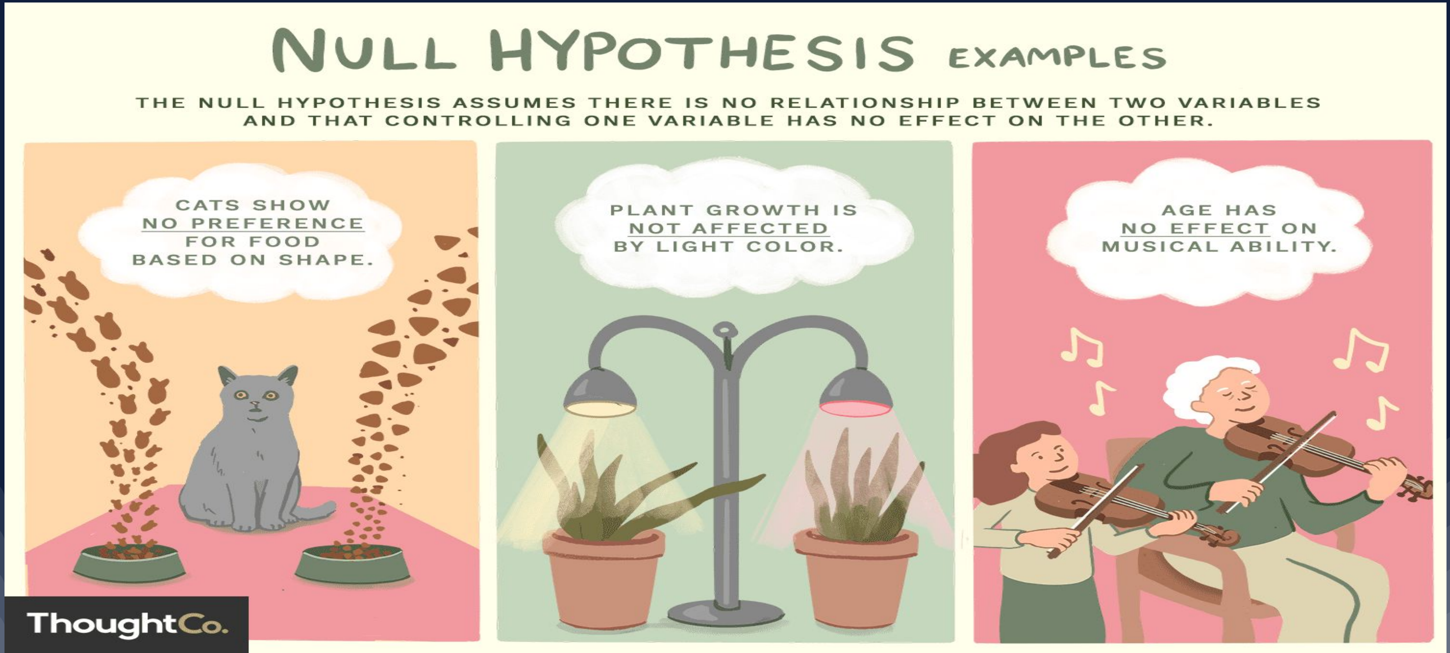
Terms to be applied

- Hypothesis testing
- The Null hypothesis
- The alternative
- Test statistic
- P-value
- Confidence interval

Definition of Terms

- **Hypothesis testing -**This is a systematic way to select samples from a group or population with the intent of making a determination about the expected behavior of the entire group.
- **The Null hypothesis -** This is a **statement** about the population that either is believed to be true or is used to put forth an argument unless it can be shown to be incorrect beyond a reasonable doubt
- **The alternative-** This is a claim about the population that is contradictory to the null hypothesis and what we conclude when we reject the null hypothesis.
- **Test statistic -** This is a value computed from sample data and is used to help us decide whether to either accept or reject the null hypothesis
- **P-value -** This is the probability of getting the observed results of a test, assuming that the null hypothesis is true

- The **null hypothesis (H_0)** is a statement in statistical hypothesis testing that assumes there is no effect, relationship, or difference in a population. It serves as the default or baseline assumption that the observed data or results occur purely by chance. The goal of hypothesis testing is to either reject or fail to reject the null hypothesis based on sample data.



The **alternative hypothesis (H_1 or Ha)** is a statement in hypothesis testing that contradicts the null hypothesis. It represents the idea that there is a real effect, relationship, or difference in the population, and it is what researchers are often hoping to support with their data. If the null hypothesis is rejected, the alternative hypothesis is considered to be more likely true

	Null	Alternate
Definition	statement that states that there is no relationship between two phenomenons under consideration or that there is no association between two groups	statement that describes that there is a relationship between two selected variables
Symol	H_0,	HIt is denoted by H_1 or H_a.
Acceptance	If the p-value is greater than the level of significance, the null hypothesis is accepted.	If the p-value is smaller than the level of significance, an alternative hypothesis is accepted
Mathematical expression	It is followed by 'equals to' sign.	followed by not equals to, 'less than' or 'greater than' sign.
Observation	believes that the results are observed as a result of chance.	believes that the results are observed as a result of some real causes
Nature	researcher tries to disprove	researcher tries to prove

Steps to determine whether to reject a null hypothesis

Step 1: state the two hypotheses so that only one can be right

Step 2 Choose the Significance Level (α) : significance level, typically set at **0.05**, defines the threshold for determining whether the results are statistically significant.

Step 3: carry out the plan and physically analyze the sample data, that involves computing the p-values, values greater than cv means rejecting the null hypothesis

Step 4: Find the P-value and compare to the significance level

If p-value ≤ α: Reject the null hypothesis (H_0). There is strong evidence that the alternative hypothesis (H_1) is true.

If p-value > α: Fail to reject the null hypothesis (H_0). There isn't enough evidence to support H_1

Random callouts

Example

Suppose a researcher wants to determine whether a new teaching method has an effect on students' test scores. The current average score is known to be 70. The researcher collects a sample of 20 students who were taught using the new method and finds that the sample's average score is 75 with a standard deviation of 10. The researcher conducts a t-test at a significance level of 0.05 to see if there is a statistically significant difference from the population mean of 70.

Errors associated with Hypothesis testing

Type 1 Error - An error that occurs when the sample results lead to rejection of the null hypothesis when in reality it is true.

Type II Error - An error that occurs when the sample results lead to acceptance of the null hypothesis when it false

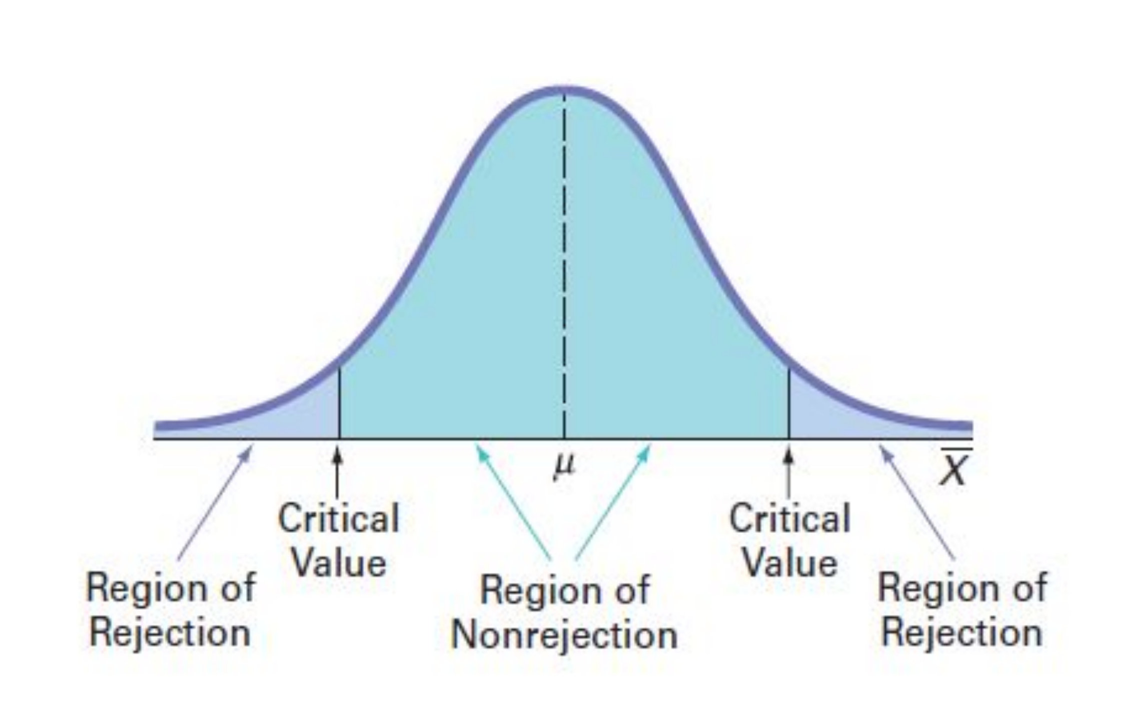
Confidence Level / Level of Confidence - It represents how certain we are in our decision. It is normally expressed as a percentage.

Power of a test - This is the probability of rejecting the null hypothesis given that it is false.

Critical area of distribution - This is the region that represents the rejection of the null hypothesis. This means that when our calculated p values fall in this region, we can decide to reject the null hypothesis

Errors associated with Hypothesis testing... Cont...

- **Confidence level:** The probability that if a poll/test/survey were repeated over and over again, the results obtained would be the same.
- **Significance level:** In a hypothesis test, the significance level, alpha, is the probability of making the wrong decision when the null hypothesis is true.



μ
$\overline{X}$
Critical Value
Critical Value
Region of Rejection
Region of Nonrejection
Region of Rejection

Type I and Type II Error

Null hypothesis is ...	True	False
Rejected	Type I error False positive Probability = α	Correct decision True positive Probability = $1 - \beta$
Not rejected	Correct decision True negative Probability = $1 - \alpha$	Type II error False negative Probability = β

Scribbr

Is a Type I or Type II error worse?

Is a Type I or Type II error worse?

For statisticians, a Type I error is usually worse. In practical terms, however, either type of error could be worse depending on your research context.

A Type I error means mistakenly going against the main statistical assumption of a null hypothesis. This may lead to new policies, practices or treatments that are inadequate or a waste of resources

How do you reduce the risk of making a Type I error?

- Choosing the right significance value. The value you set at the beginning of your study to assess the statistical probability of obtaining your results (p value)
- The significance level is usually set at 0.05 or 5%. This means that your results only have a 5% chance of occurring, or less, if the null hypothesis is actually true.
- To reduce the Type I error probability, you can set a lower significance level.

How do you reduce the risk of making a Type II error?

The risk of making a Type II error is inversely related to the statistical power of a test. Power is the extent to which a test can correctly detect a real effect when there is one.

- Increase Sample Size
- Increase the Significance Level (α)
- Use a More Sensitive Test
- Increase the Effect Size
- Use One-Tailed Tests (When Appropriate)

Random callouts

Example

Scenario:

Suppose you are a researcher testing the effectiveness of a new drug that claims to lower blood pressure. Your hypothesis test involves:

State the Null and Alternative Hypothesis

Example

Scenario:

Suppose you are a researcher testing the effectiveness of a new drug that claims to lower blood pressure. Your hypothesis test involves:

- **Null Hypothesis (H_0):** The new drug has no effect on blood pressure (it is no different from a placebo).
- **Alternative Hypothesis (H_1):** The new drug lowers blood pressure (it has a significant effect).
-

Key Differences between Type I & II Errors

- Type I error is the same as a false alarm or false positive while Type II error is also referred to as false negative.

- A Type I error is represented by α while a Type II error is represented by β

- The level of significance determines the possibility of a type I error while type II error is the possibility of deducting the power of the test from 1

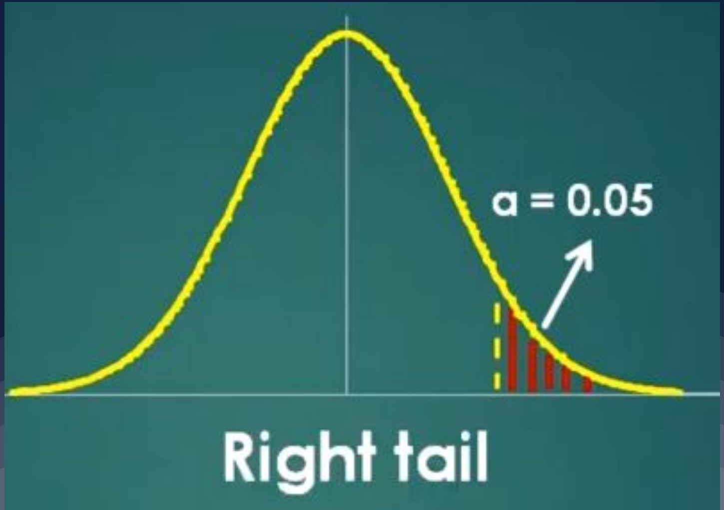
- In statistical hypothesis testing, a type I error is caused by disapproving a null hypothesis that is otherwise correct while in contrast, Type II error occurs when the null hypothesis is not rejected even though it is not true.
- Type I error occurs when you reject the null hypothesis, in contrast, Type II error occurs when you accept an incorrect outcome of a false hypothesis

One Tail and Two Tail Test

A one-tail test is a statistical test where the critical region of the distribution is one-sided

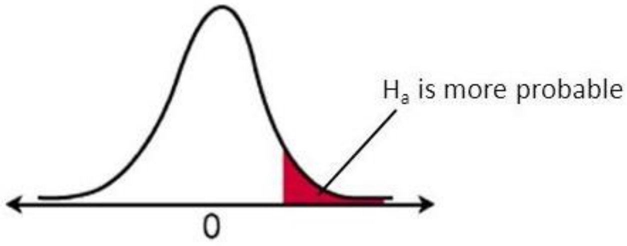
It can either greater than or less than a certain value, but it cannot be both

A one-tail test can be further classified as either a right-tail test or a left tail test



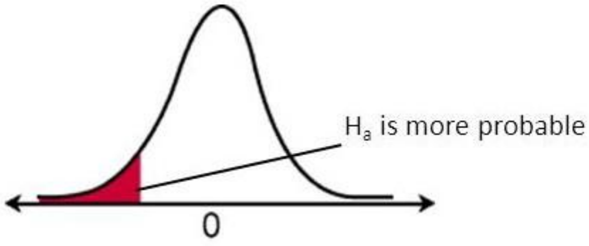
One Tail and Two Tail Test

- If the sample being tested falls in either of the two regions then the alternative hypothesis is accepted.
- Example
- we may wish to compare the mean of a sample to a given value *x* using a t-test. Our null hypothesis is that the mean is equal to *x*. A two-tailed test will test both if the mean is significantly greater than *x* and if the mean significantly less than *x*.
- The mean is considered significantly different from *x* if the test statistic is in the top 2.5% or bottom 2.5% of its probability distribution, resulting in a p-value less than 0.05.



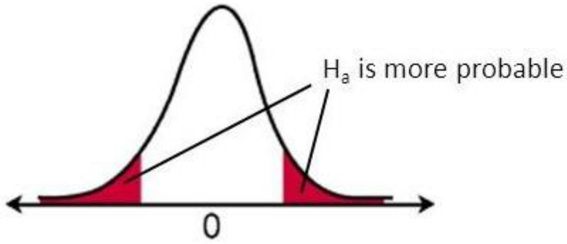
Right-tail test

H_a: $\mu >$ value



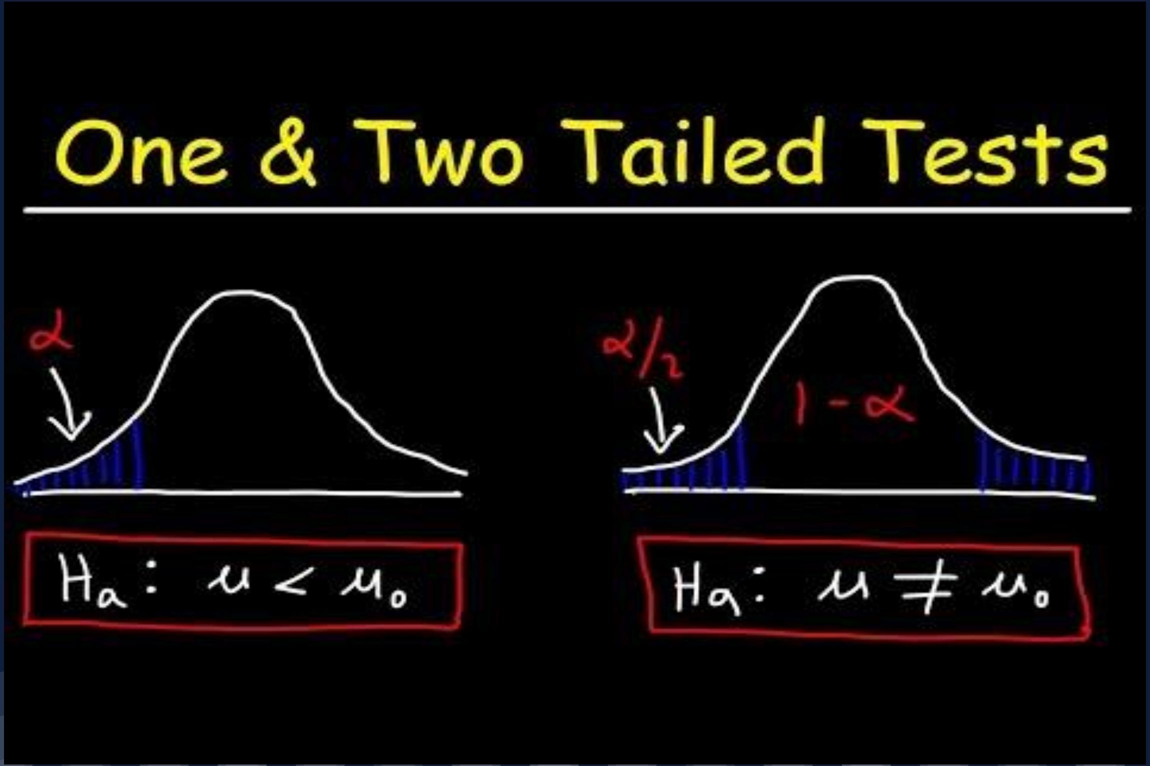
Left-tail test

H_a: $\mu <$ value



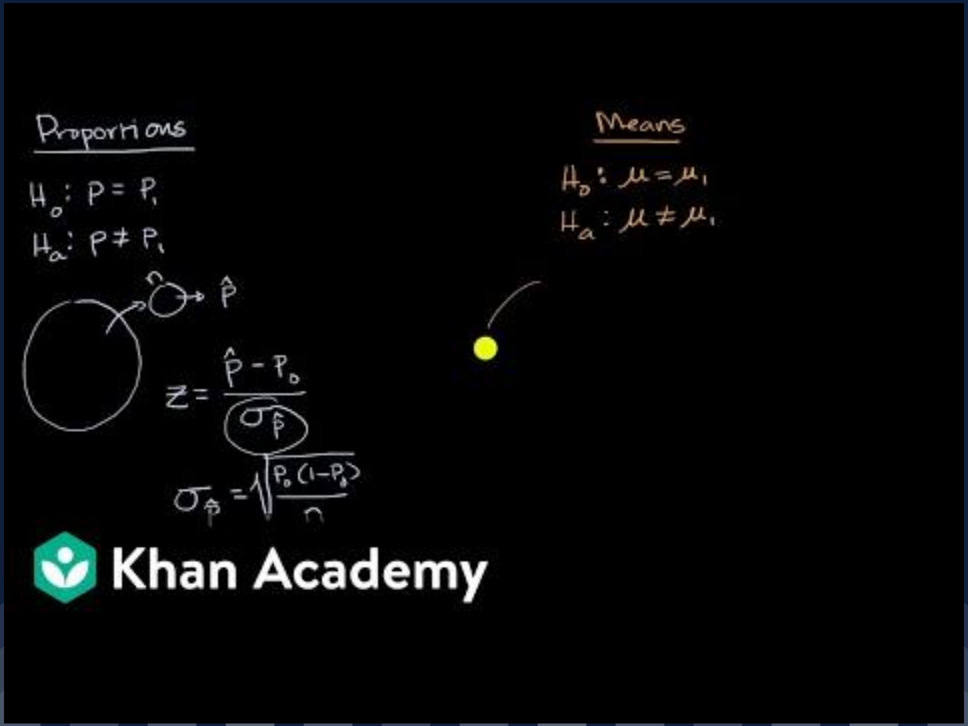
Two-tail test

H_a: $\mu \neq$ value


One & Two Tailed Tests
α
α/2
1 - α
Ha : u < u₀
Ha : u ≠ u₀

Test Statistic

- A test Statistic is a random variable that is calculated from sample data and used in a hypothesis test. Either to reject or accept the null hypothesis.



P-value

A **p-value** is a number that tells us how likely it is to get the results we observed in a study, assuming the **null hypothesis** (no effect or difference) is true. It helps us decide if our findings are statistically significant or just due to chance.

- **Small p-value (usually less than 0.05):** There is strong evidence that what we observed is real, and we should reject the null hypothesis.
- **Large p-value (greater than 0.05):** There is not enough evidence to say the result is significant, so we do not reject the null hypothesis.

Example of P-value

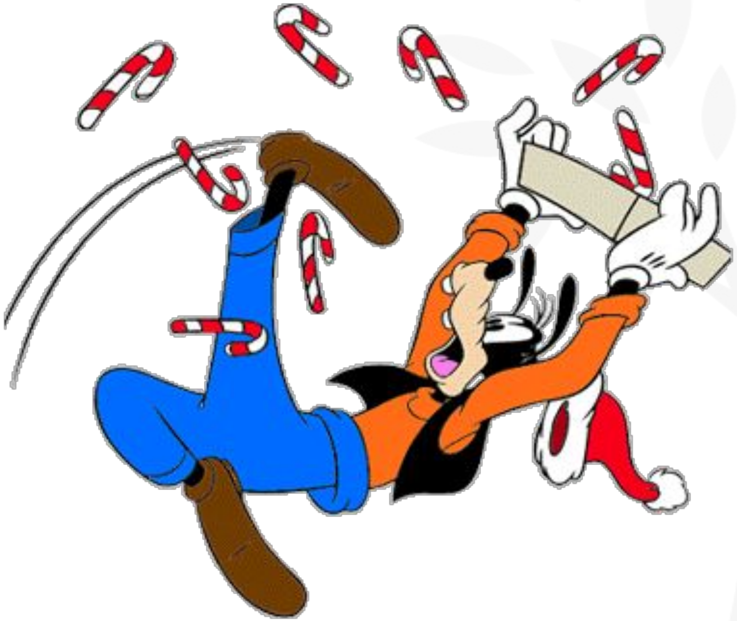
Suppose you want to test if a new teaching method improves student test scores compared to the old method.

- **Null hypothesis (H_0):** There is no difference between the new and old teaching methods.
- **Alternative hypothesis (H_1):** The new teaching method improves scores.

You conduct a test and find a **p-value of 0.02**.

- **Interpretation:** Since the p-value (0.02) is less than 0.05, you can reject the null hypothesis and conclude that the new teaching method likely improves scores. The result is statistically significant, meaning it's unlikely to be due to chance.

Any Stumbles ?



Thanks!

Have an amazing week ahead. Stay safe.

Sam

We value you!