

Inferential Analysis

Part 4

Introduction to Hypothesis

A hypothesis test evaluates two statements about a population. The statements are mutually exclusive. The test concludes which statement best reflects the sample data. A hypothesis test helps us determine the statistical significance of a finding.

Null Hypothesis (H_0) : Null hypothesis is the hypothesis which is tested for the possible rejection under the assumption that is true

Alternative Hypothesis (H_a) : Alternative hypothesis is the logical opposite of the null hypothesis

Null Hypothesis

H_0

A statement about a population parameter.

We test the likelihood of this statement being true in order to decide whether to accept or reject our alternative hypothesis.

Can include =, ≤, or ≥ sign.



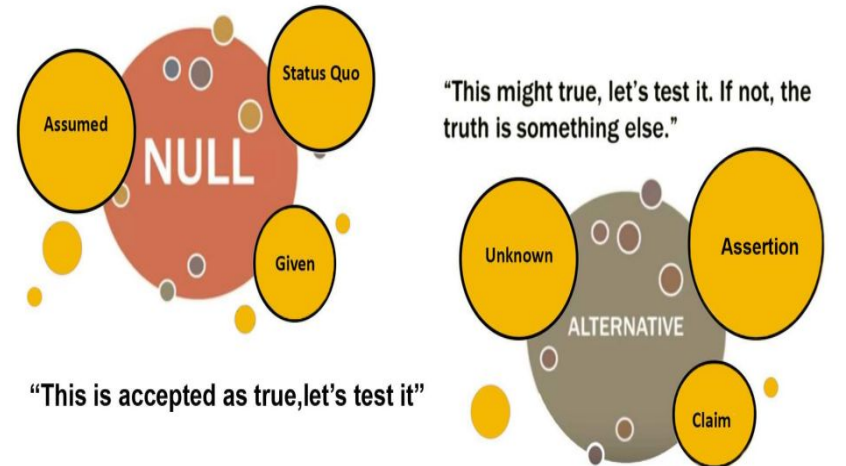
Alternative Hypothesis

H_a

A statement that directly contradicts the null hypothesis.

We determine whether or not to accept or reject this statement based on the likelihood of the null (opposite) hypothesis being true.

Can include a ≠, >, or < sign.



Significance Level (alpha : α): Criterion used for rejecting the null hypothesis. **Red line in figure**

P-value: The probability of outcomes more extreme than the observed outcome, assuming the null is true. **Area to the left of the blue line**

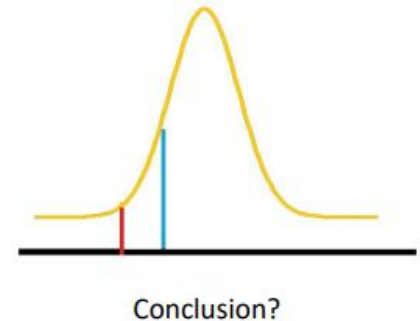
Conclusion: If p-value is less than significance level reject null hypothesis.

P-value:

- The p-value is the level of marginal significance within a statistical hypothesis test representing the probability of the occurrence of a given event.
- p-value also known as rejection region
- The p-value is used as an alternative to rejection points to provide the smallest level of significance at which the null hypothesis would be rejected.
- A smaller p-value means that there is stronger evidence in favour of the alternative hypothesis.
- The rejection region is found by using alpha to find a critical value;
- The rejection region is the area that is more extreme than the critical value.

What if p-value > Significance level?

Basic Setup in Hypothesis Test	When sales is 19,800	When sales is 18,000
Null Hypothesis	Reduction in daily sales	Reduction in daily sales
Alternate Hypothesis	No reduction in daily sales	No reduction in daily sales
Test Distribution	Normal Distribution	Normal Distribution
Significance Level	5%	5%
p-value	0.401	0.012
Conclusion	Accept Null Hypothesis	Reject Null Hypothesis

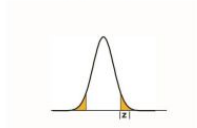


Types of Alternative Hypothesis

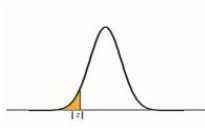
There are three types of alternative hypotheses:

- ✓ Two-sided test – When the population parameter is not equal to a certain value.
 $H_a: p \neq p_0$, or $H_a: \mu \neq \mu_0$
- ✓ Lower or Left-tailed test - When the population parameter is less than a certain value.
 $H_a: p < p_0$, or $H_a: \mu < \mu_0$
- ✓ Upper or Right-tailed test – When the population parameter is greater than a certain value.
 $H_a: p > p_0$, or $H_a: \mu > \mu_0$
 Where, p = Proportion
 μ = Mean

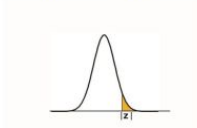
Two Tail test



Left tail test



Right tail test



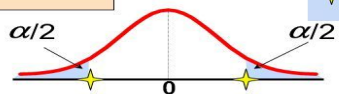
Level of Significance and the Rejection Region

Level of significance = α

$$H_0: \mu = 3$$

$$H_1: \mu \neq 3$$

Two-tail test



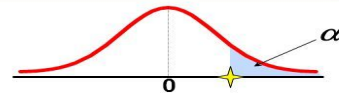
★ Represents critical value

Rejection region is shaded

$$H_0: \mu \leq 3$$

$$H_1: \mu > 3$$

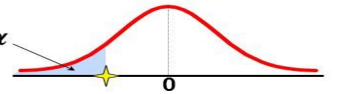
Upper-tail test



$$H_0: \mu \geq 3$$

$$H_1: \mu < 3$$

Lower-tail test

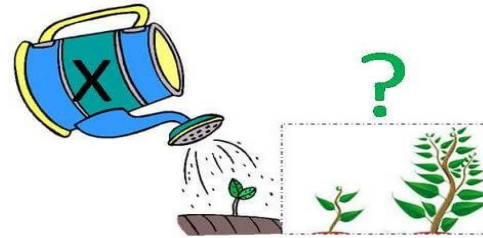


Effect of Bio-fertilizer 'x' on Plant growth

www.majordifferences.com

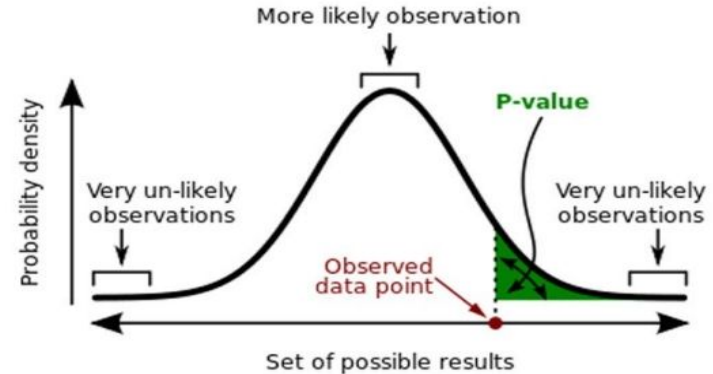
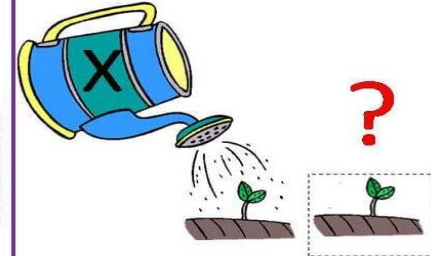
Alternative Hypothesis

H_1 : Application of bio-fertilizer 'x' increase plant growth.



Null Hypothesis

H_0 : Application of bio-fertilizer 'x' do not increase plant growth.



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Types of Errors in Hypothesis Testing

In Hypothesis testing, there are two kinds of errors that can be made in significance testing

- The risks of these two errors are inversely related and determined by the level of significance and the power of the test. Therefore, you should determine which error has more severe consequences for your situation before you define their risks.

Type 1 Error : This error occurs if we reject the null hypothesis H_0 (in favor of alternative hypothesis (H_a)) when the null hypothesis (H_0) is true , **We denote $\alpha = P(\text{Type I Error})$.**

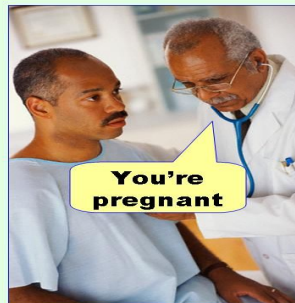
Type 2 Error : This error occurs if we fail to reject the null hypothesis H_0 when the alternative hypothesis H_a is true.

We denote $\beta = P(\text{Type II Error})$.

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

Type I error
(false positive)



Type II error
(false negative)



Choosing a Statistical Test

Next slide Image:

Link: <https://www.intro2r.info/unit3/which-test.html>

Image source: www.tutor2u.net

