AUA CS 108, Statistics, Fall 2019

Lecture 36

Michael Poghosyan
YSU, AUA
michael@ysu.am, mpoghosyan@aua.am

15 Nov 2019

Contents

- Hypothesis Testing

Last Lecture ReCap

- Give the Hypo Testing Framework.

Last Lecture ReCap

- Give the Hypo Testing Framework.
- Give the usual types of Hypotheses

Moral, and Choosing Null Hypotheses

Moral: In Hypothesis tesing, if we have enough evidence from Data against $\mathcal{H}_0$, we Reject it, otherwise, we say that we do not have enough evidence to Reject $\mathcal{H}_0$, so we Fail to Reject it, and keep believing in it.

Moral, and Choosing Null Hypotheses

Moral: In Hypothesis tesing, if we have enough evidence from Data against $\mathcal{H}_0$, we Reject it, otherwise, we say that we do not have enough evidence to Reject $\mathcal{H}_0$, so we Fail to Reject it, and keep believing in it.

One is using this general idea to choose the Null and Alternative Hypotheses:

Moral, and Choosing Null Hypotheses

Moral: In Hypothesis tesing, if we have enough evidence from Data against $\mathcal{H}_0$, we Reject it, otherwise, we say that we do not have enough evidence to Reject $\mathcal{H}_0$, so we Fail to Reject it, and keep believing in it.

One is using this general idea to choose the Null and Alternative Hypotheses: **we will keep believing in Null, if the Data will not show strong evidence against**.

Examples

Let us consider few examples:

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if our new marketing campaign (new method of advertising) is doing better. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if our new marketing campaign (new method of advertising) is doing better. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if our new marketing campaign (new method of advertising) is doing better. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if the new developed design of our webpage/app is more attractive than the recent one. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if our new marketing campaign (new method of advertising) is doing better. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if the new developed design of our webpage/app is more attractive than the recent one. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?

Examples

Let us consider few examples:

- We want to test if the email message is spam. How to choose the Null and Alternative? How to Test it?
- We want to see if the person has some illness. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if a new drug, developed by some company, works better than the drug people are using now. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if our new marketing campaign (new method of advertising) is doing better. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?
- We want to see if the new developed design of our webpage/app is more attractive than the recent one. How to choose $\mathcal{H}_0$ and $\mathcal{H}_1$? How to Test it?

This is an example of **A/B Testing**.

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative?

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?
- Assume we are generating some numbers, and we want to see if they are random or not. How to choose the Null and Alternative?

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?
- Assume we are generating some numbers, and we want to see if they are random or not. How to choose the Null and Alternative? How to Test it?

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?
- Assume we are generating some numbers, and we want to see if they are random or not. How to choose the Null and Alternative? How to Test it? See, for example, Diehard Tests.

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?
- Assume we are generating some numbers, and we want to see if they are random or not. How to choose the Null and Alternative? How to Test it? See, for example, Diehard Tests.
- We want to see if boy and girl child births rates are the same in our country. How to choose the Null and Alternative?

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?
- Assume we are generating some numbers, and we want to see if they are random or not. How to choose the Null and Alternative? How to Test it? See, for example, Diehard Tests.
- We want to see if boy and girl child births rates are the same in our country. How to choose the Null and Alternative? How to Test it?
- We want to see if smoking causes lung cancer. How to choose the Null and Alternative?

Examples

- Assume we have some Observations, and we want to see if that Observations come from some particular Distribution, say, they are from the Normal Distribution. How to choose the Null and Alternative? How to Test it?
- Assume we are generating some numbers, and we want to see if they are random or not. How to choose the Null and Alternative? How to Test it? See, for example, Diehard Tests.
- We want to see if boy and girl child births rates are the same in our country. How to choose the Null and Alternative? How to Test it?
- We want to see if smoking causes lung cancer. How to choose the Null and Alternative? How to Test it?

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	
Reject $\mathcal{H}_0$	

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	
Reject $\mathcal{H}_0$	
Do Not Reject $\mathcal{H}_0$	

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	$\mathcal{H}_0$ is True
Reject $\mathcal{H}_0$	
Do Not Reject $\mathcal{H}_0$	

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	$\mathcal{H}_0$ is True	$\mathcal{H}_0$ is False (i.e., $\mathcal{H}_1$ is True)
Reject $\mathcal{H}_0$		
Do Not Reject $\mathcal{H}_0$		

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	$\mathcal{H}_0$ is True	$\mathcal{H}_0$ is False (i.e., $\mathcal{H}_1$ is True)
Reject $\mathcal{H}_0$		Correct Decision (True Negative)
Do Not Reject $\mathcal{H}_0$		

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	$\mathcal{H}_0$ is True	$\mathcal{H}_0$ is False (i.e., $\mathcal{H}_1$ is True)
Reject $\mathcal{H}_0$		Correct Decision (True Negative)
Do Not Reject $\mathcal{H}_0$	Correct Decision (True Positive)	

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

Then the following cases can happen:

Test Decision \ Reality	$\mathcal{H}_0$ is True	$\mathcal{H}_0$ is False (i.e., $\mathcal{H}_1$ is True)
Reject $\mathcal{H}_0$	**Type I Error (False Positive)**	Correct Decision (True Negative)
Do Not Reject $\mathcal{H}_0$	Correct Decision (True Positive)	

Type I and II errors (aka Confusion Matrix)

Assume we are Testing the Hypothesis

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1.$$

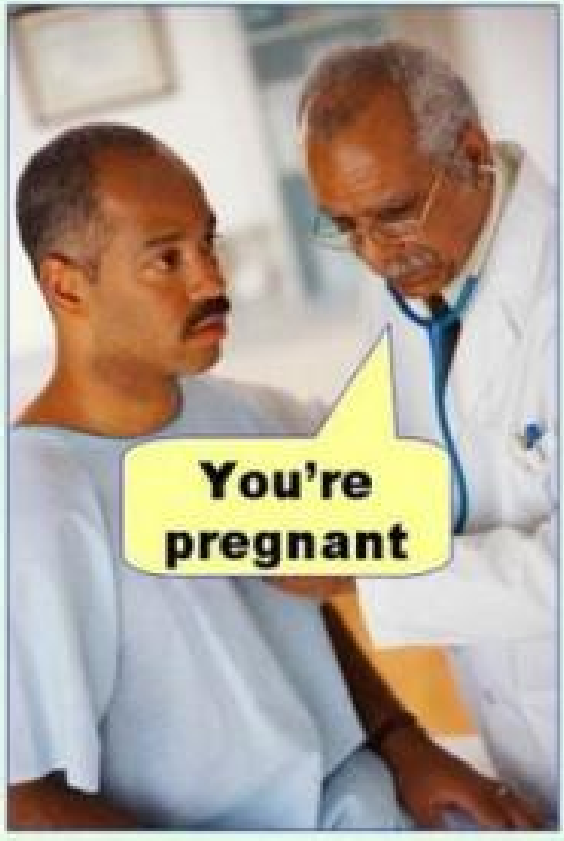
Then the following cases can happen:

Test Decision \ Reality	$\mathcal{H}_0$ is True	$\mathcal{H}_0$ is False (i.e., $\mathcal{H}_1$ is True)
Reject $\mathcal{H}_0$	**Type I Error (False Positive)**	Correct Decision (True Negative)
Do Not Reject $\mathcal{H}_0$	Correct Decision (True Positive)	**Type II Error (False Negative)**

Can you guess the Null Hypo?

Can you guess the Null Hypo?

Type I error
(false positive)



Type II error
(false negative)