

# Malaria Molecular Surveillance Study Design Workshop

## Module 3: Hypothesis testing

## Null hypothesis testing

---



Sometimes we are simply trying to estimate something, e.g. prevalence. **We have seen how to perform sample size calculation based on precision arguments.**

In other cases, we have a specific question that we want to answer. This questions may be phrased as a **null hypothesis test**.

---

## Null hypothesis testing



Sometimes we are simply trying to estimate something, e.g. prevalence. **We have seen how to perform sample size calculation based on precision arguments.**

In other cases, we have a specific question that we want to answer. This questions may be phrased as a **null hypothesis test**.

---

A null hypothesis ( $H_0$ ) is a statement of **no effect or difference** between groups. This is often a statement that nothing interesting is happening\*

*Rather than trying to prove there is an effect, in null hypothesis testing we try to disprove that there is no effect.*

\* Sometimes it can be very interesting if the null hypothesis is true

## Null hypothesis testing



- Q: Has prevalence increased over the last 5 years?
- Q: Are certain genetic variants associated with gender, or occupation?
- Q: Does vaccine efficacy vary based on genetic markers?

## Null hypothesis testing



- Q: Has prevalence increased over the last 5 years?  
 $H_0$ : Prevalence has remained the same over the last 5 years.
- Q: Are certain genetic variants associated with gender, or occupation?
- Q: Does vaccine efficacy vary based on genetic markers?

## Null hypothesis testing



- Q: Has prevalence increased over the last 5 years?  
 $H_0$ : Prevalence has remained the same over the last 5 years.
- Q: Are certain genetic variants associated with gender, or occupation?  
 $H_0$ : There is no association between genetic variant and gender or occupation.
- Q: Does vaccine efficacy vary based on genetic markers?

## Null hypothesis testing



- Q: Has prevalence increased over the last 5 years?  
 $H_0$ : Prevalence has remained the same over the last 5 years.
- Q: Are certain genetic variants associated with gender, or occupation?  
 $H_0$ : There is no association between genetic variant and gender or occupation.
- Q: Does vaccine efficacy vary based on genetic markers?  
 $H_0$ : Vaccine efficacy is the same irrespective of genetic markers.

## Null hypothesis testing



Each test has a **test statistic**

**One-sample z-test for proportions:** tests prevalence against a known value

## Null hypothesis testing



Each test has a **test statistic**

**One-sample z-test for proportions:** tests prevalence against a known value

$H_0$ : The population prevalence equals  $p_0$

## Null hypothesis testing



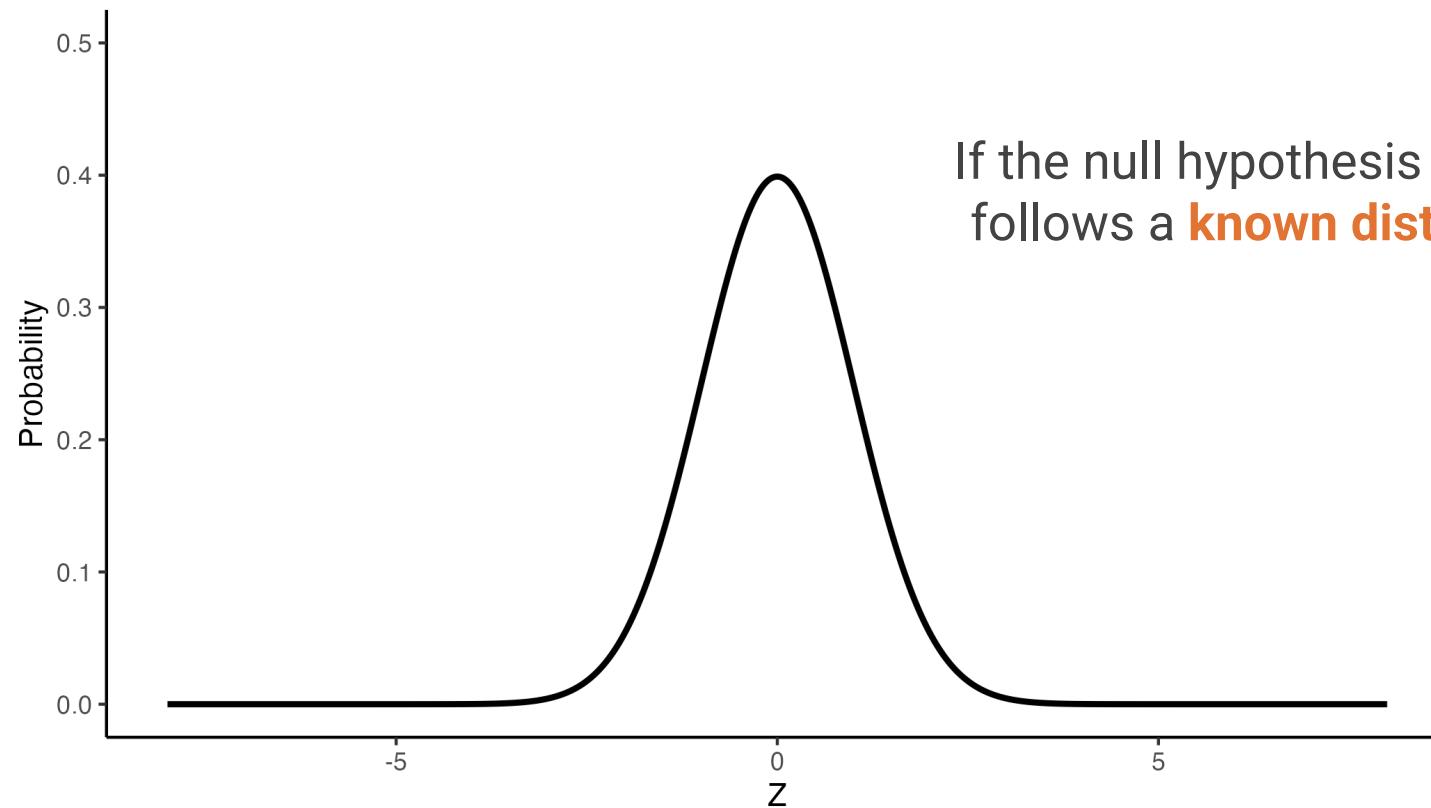
Each test has a **test statistic**

**One-sample z-test for proportions:** tests prevalence against a known value

$H_0$ : The population prevalence equals  $p_0$

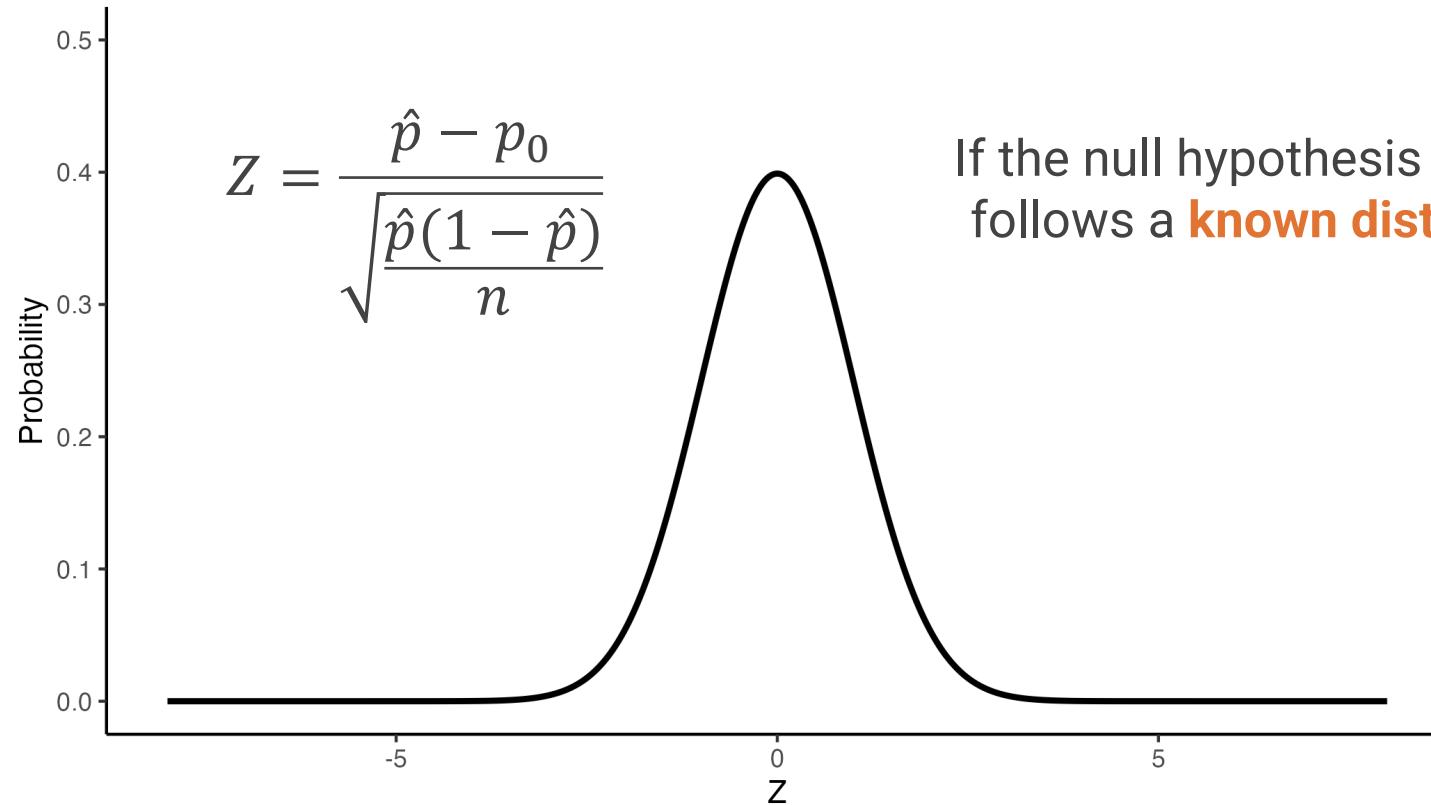
$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}}$$

## Null hypothesis testing

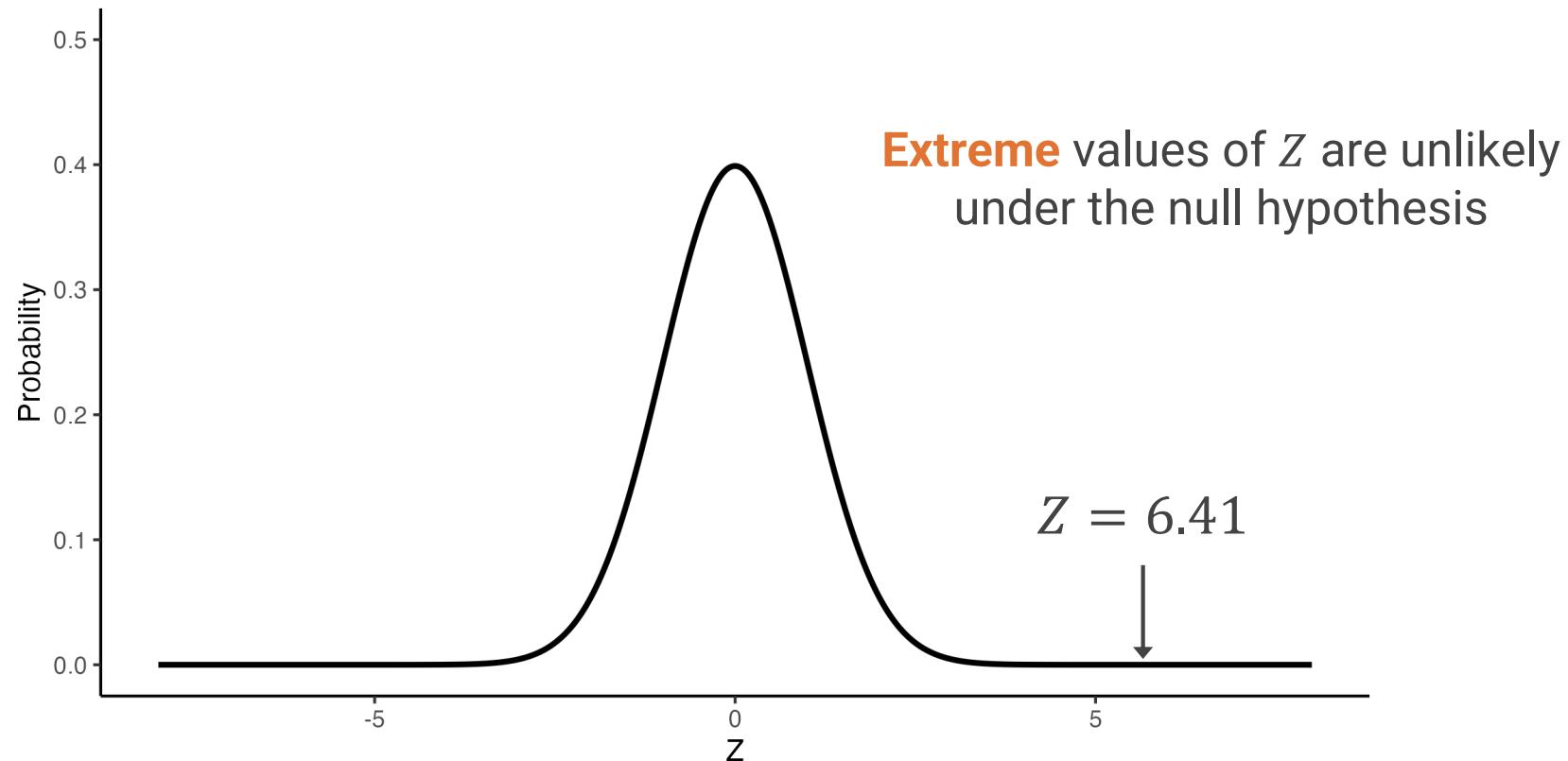


If the null hypothesis is true,  $Z$  follows a **known distribution**

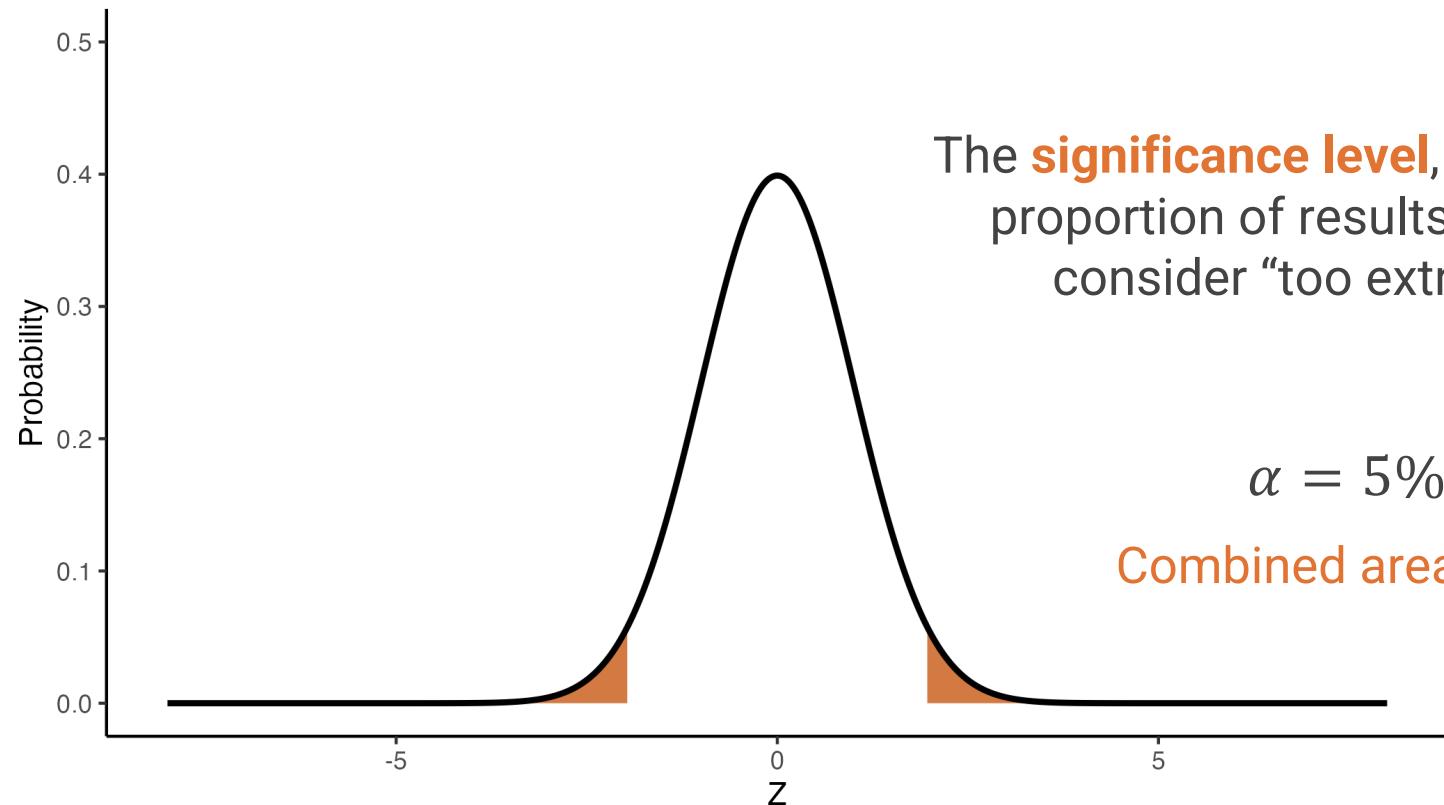
## Null hypothesis testing



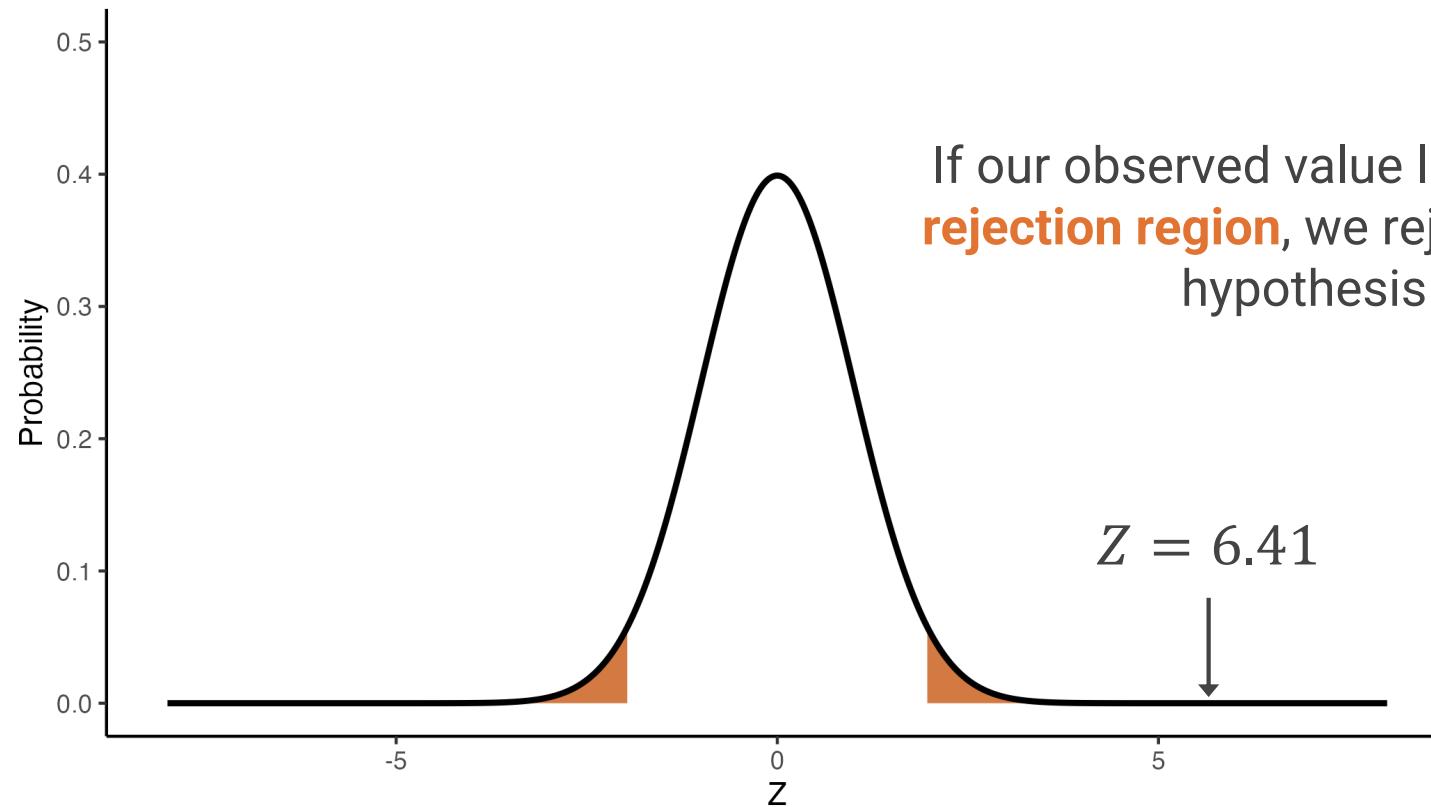
## Null hypothesis testing



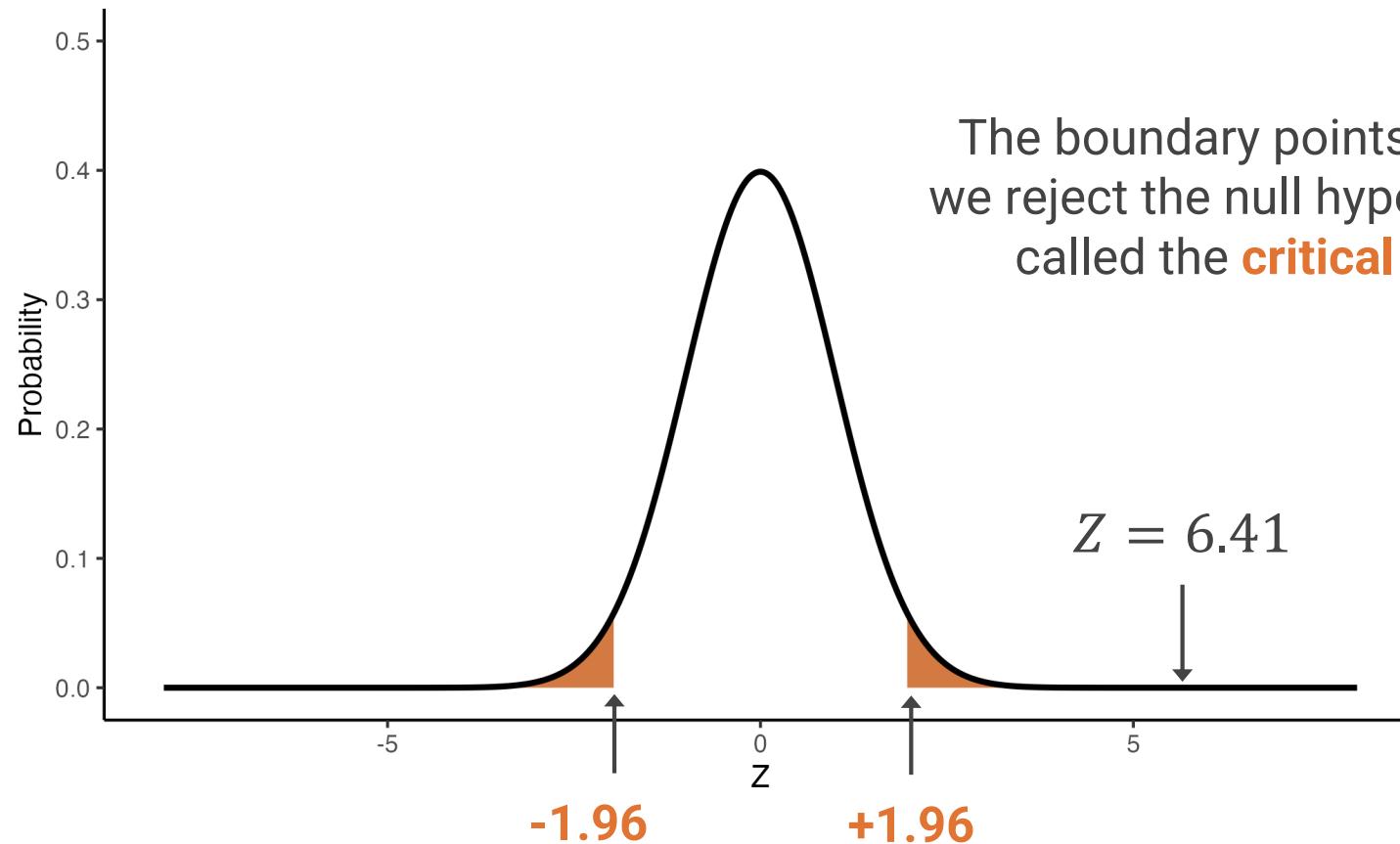
## Null hypothesis testing



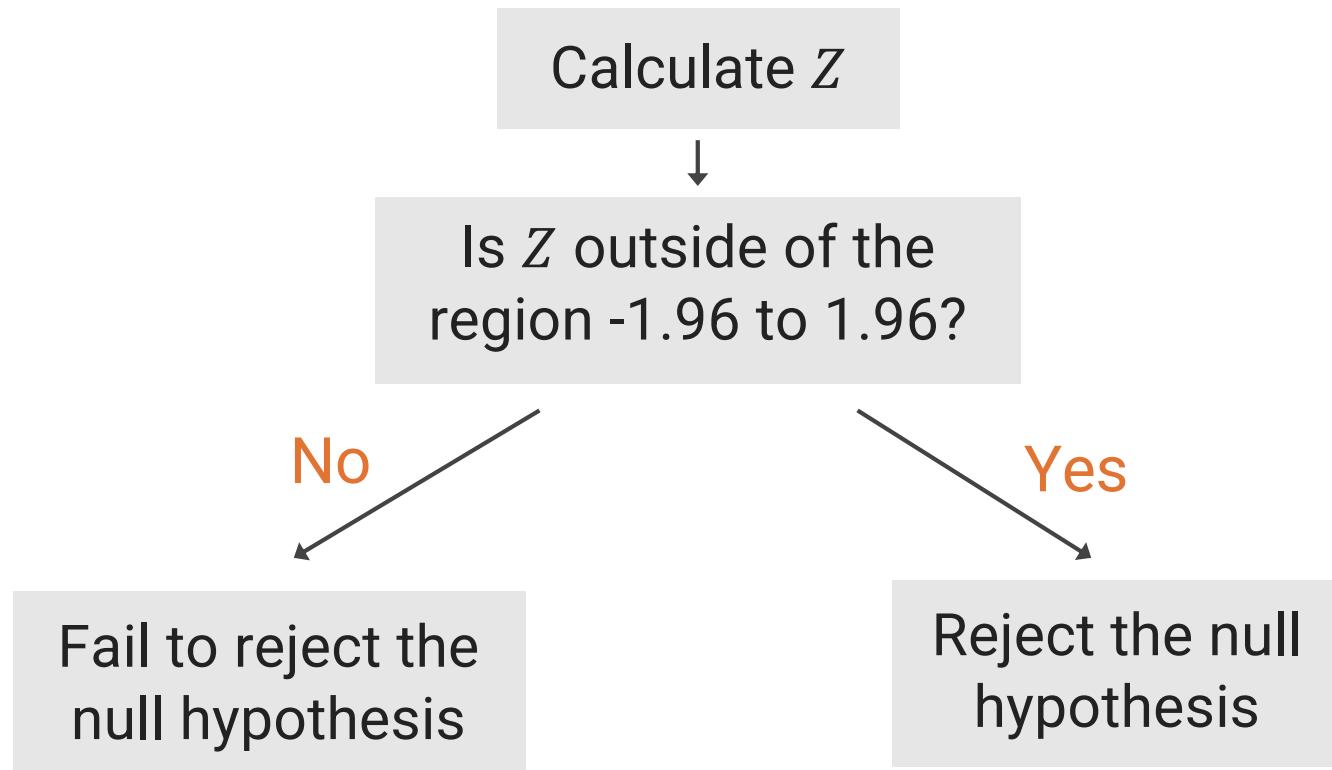
## Null hypothesis testing



## Null hypothesis testing



## Null hypothesis testing



## Null hypothesis testing



		Conclusion about $H_0$	
		Fail to reject	Reject
Truth about $H_0$	True	True negative	False positive
	False		

## Null hypothesis testing



		Conclusion about $H_0$	
		Fail to reject	Reject
Truth about $H_0$	True	True negative $1 - \alpha$	False positive $\alpha$
	False		

## Null hypothesis testing



		Conclusion about $H_0$	
		Fail to reject	Reject
Truth about $H_0$	True	True negative $1 - \alpha$	False positive $\alpha$
	False		

$\alpha$  sets the **false positive rate** of a test. Using  $\alpha$  we can control how often we incorrectly conclude that there is a real effect when there is none.

## Null hypothesis testing



		Conclusion about $H_0$	
		Fail to reject	Reject
Truth about $H_0$	True	True negative $1 - \alpha$	False positive $\alpha$
	False	What about this!? (next lecture!)	

$\alpha$  sets the **false positive rate** of a test. Using  $\alpha$  we can control how often we incorrectly conclude that there is a real effect when there is none.

## Summary



- We can ask questions using **null hypothesis tests**
- A null hypothesis is a statement of **no effect/difference** between groups
- The significance level  $\alpha$  controls the **false-positive rate**