Malaria Molecular Surveillance Study Design Workshop

Module 3: Hypothesis testing and power

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 tests**.

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

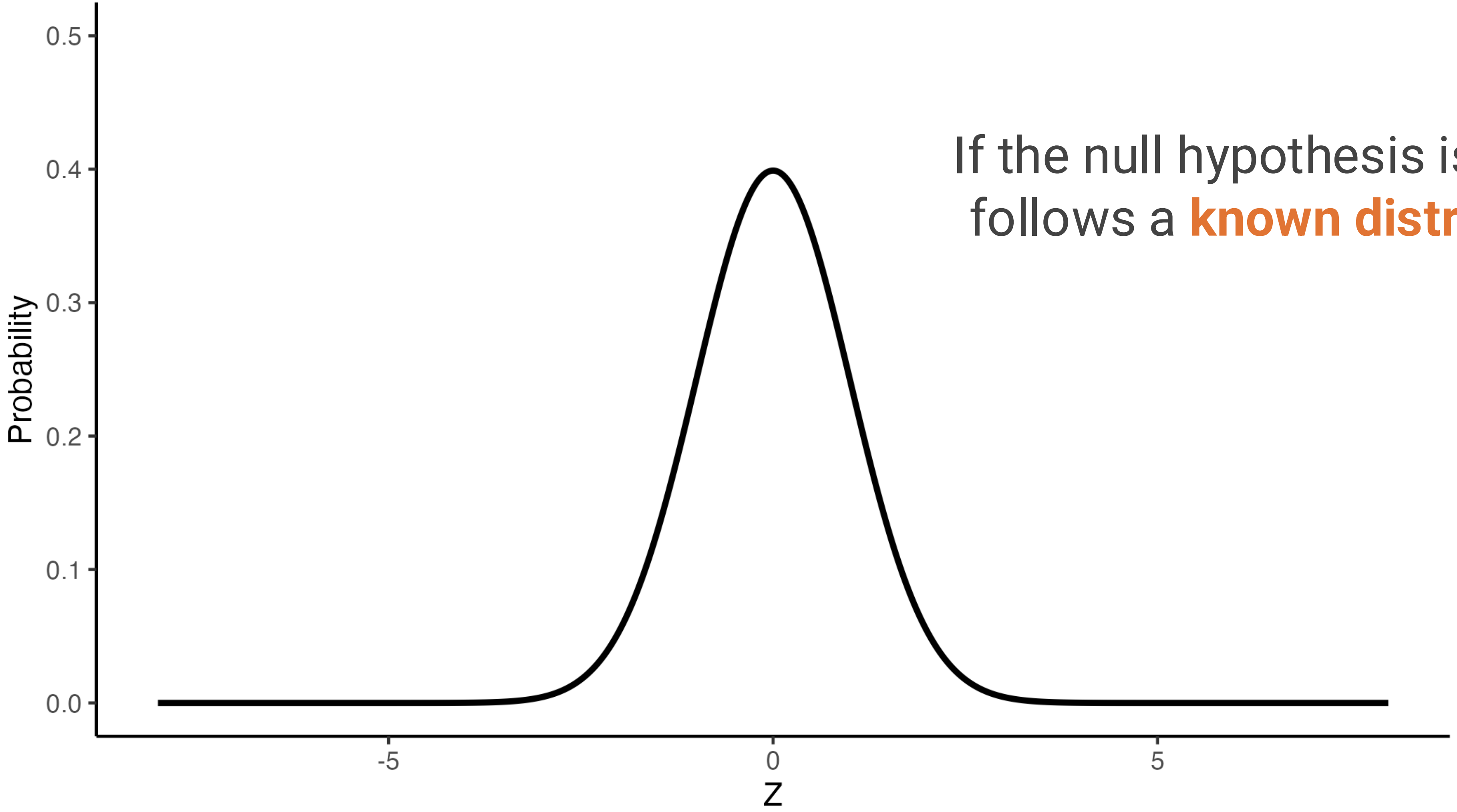
Each test has a **test statistic**

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

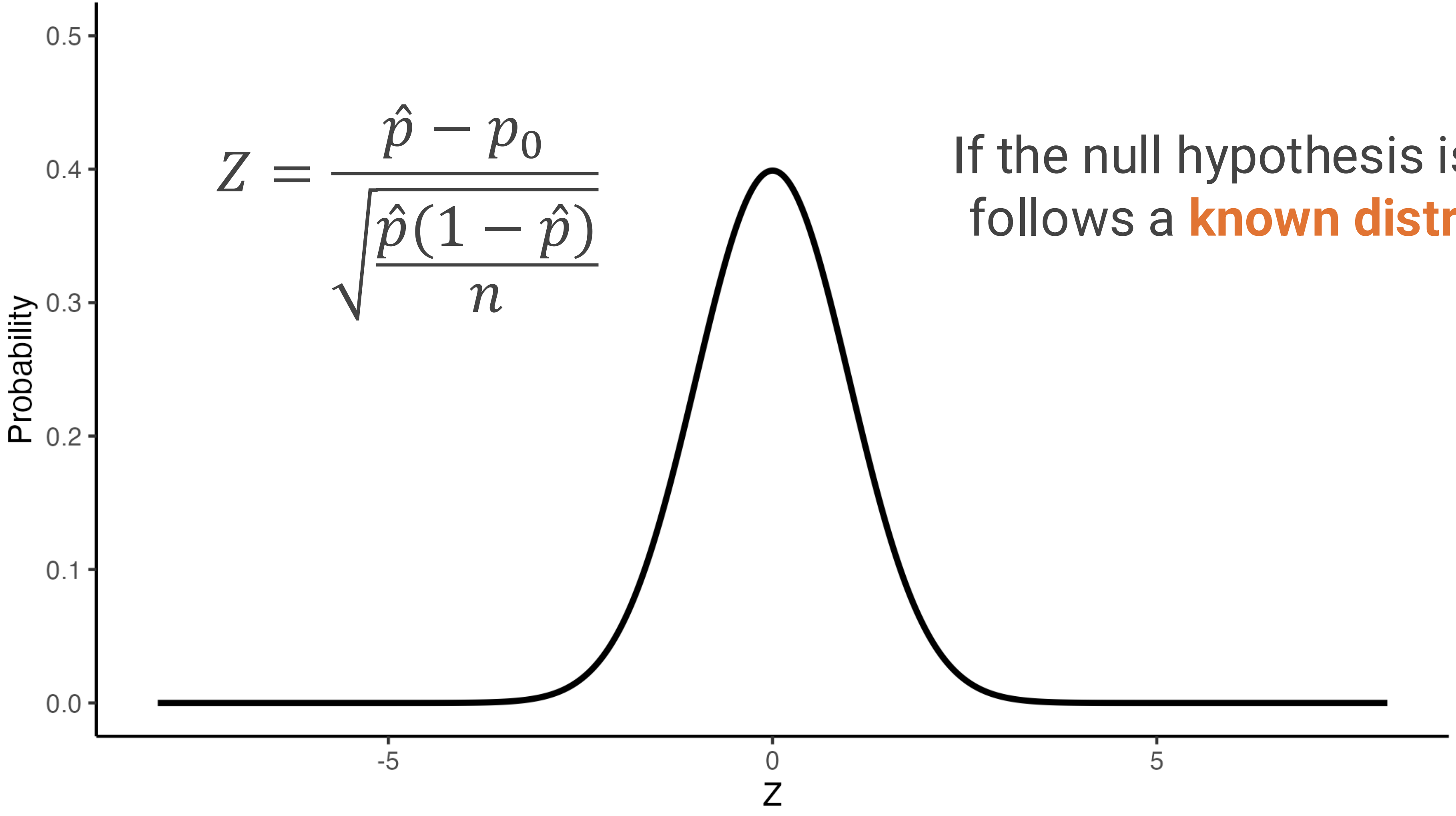


If the null hypothesis is true, Z follows a **known distribution**

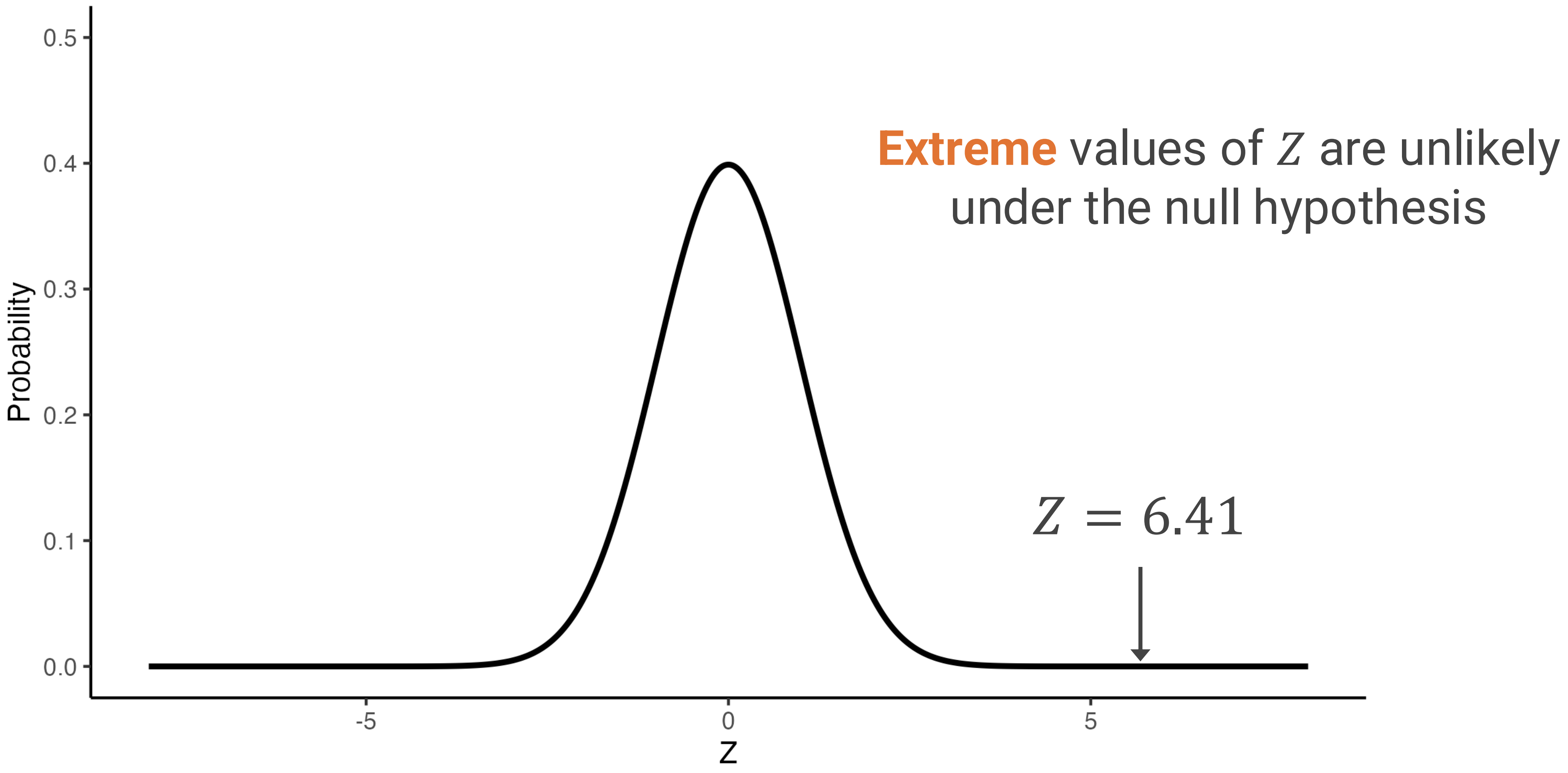
Null hypothesis testing

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

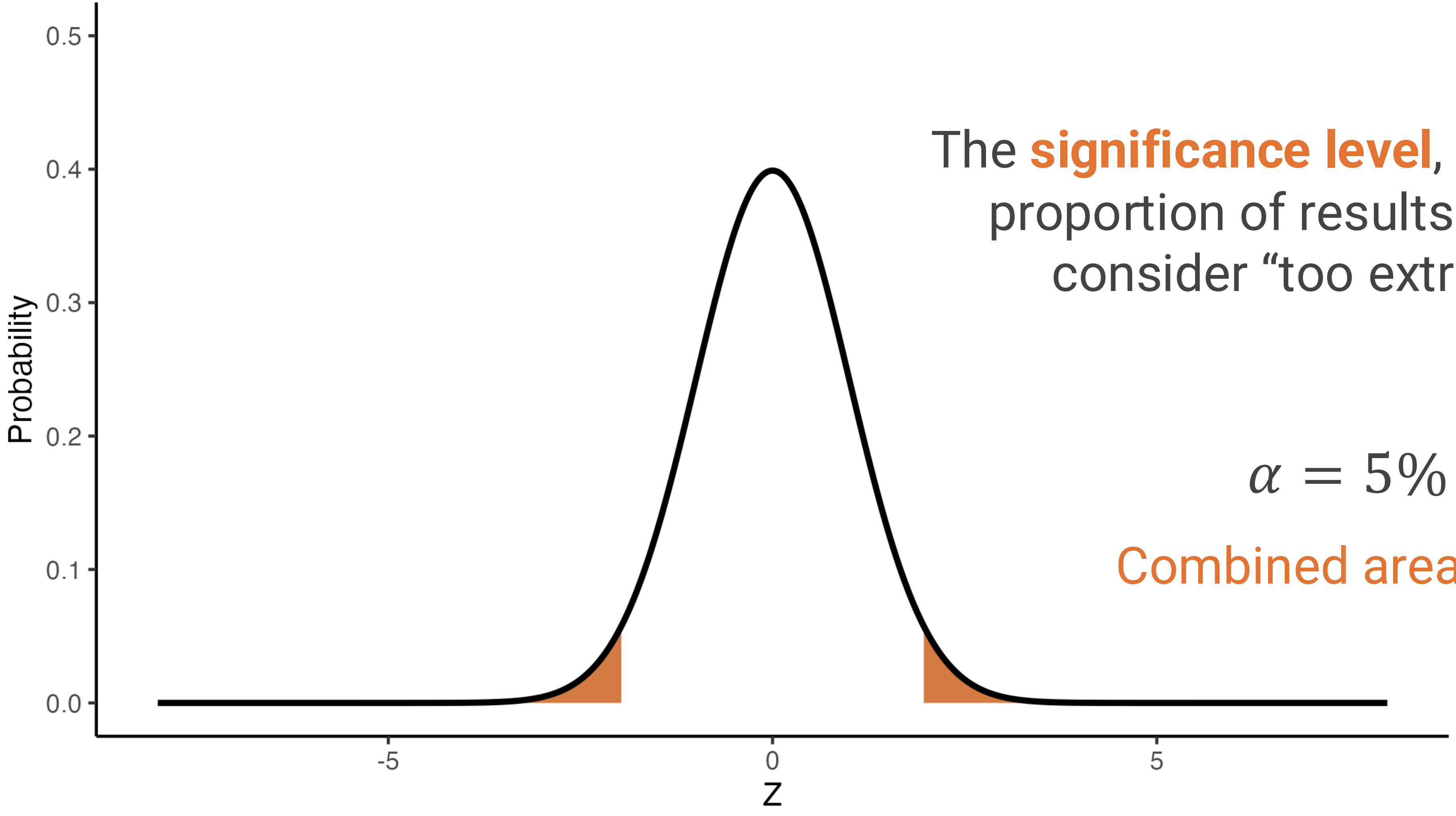
If the null hypothesis is true, Z follows a **known distribution**



Null hypothesis testing



Null hypothesis testing

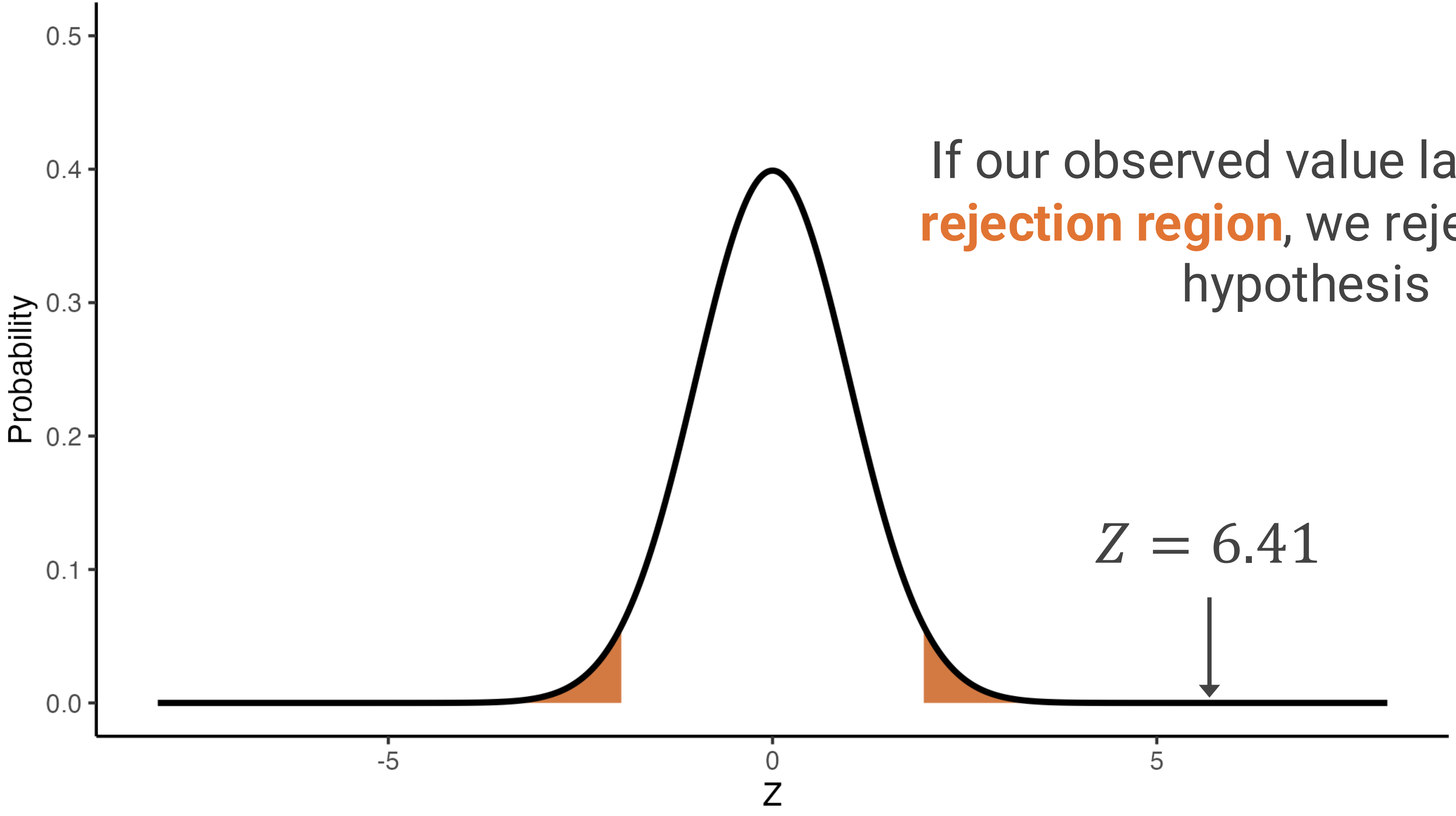


The **significance level**, α, sets the proportion of results that we consider "too extreme"

$$\alpha = 5\%$$

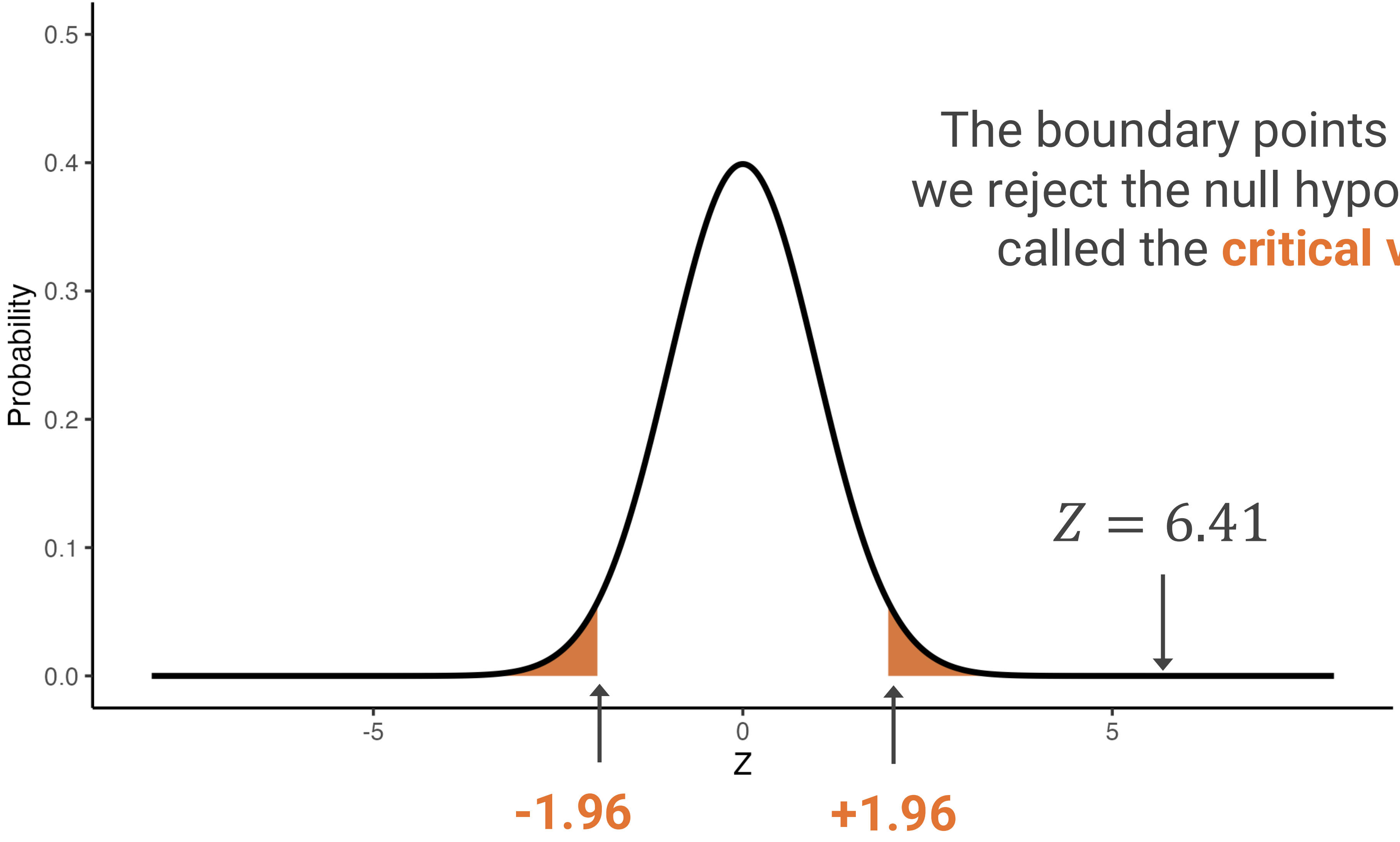
Combined area = 5%

Null hypothesis testing



If our observed value lands in the **rejection region**, we reject the null hypothesis

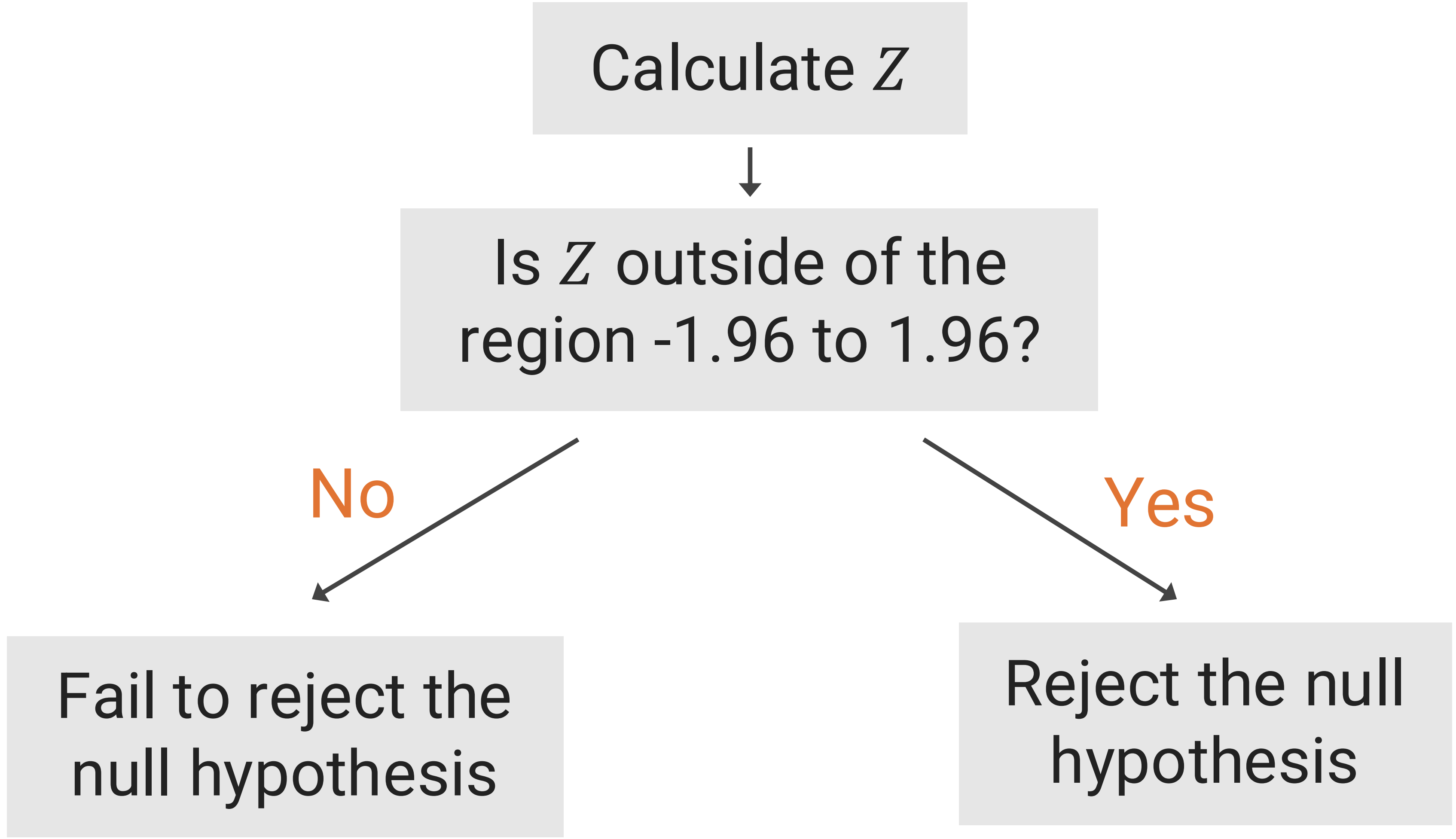
Null hypothesis testing



The boundary points at which we reject the null hypothesis are called the **critical values**

$$Z = 6.41$$

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 α
	False		

Null hypothesis testing

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

α sets the **false positive rate** of a test. Using α 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 α
	False	**What about this!?**	

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

Power

In power analysis, we also specify an **alternative hypothesis**

H_0: The population prevalence equals p_0

H_1: The population prevalence equals p, which is different from p_0

Power

In power analysis, we also specify an **alternative hypothesis**

H_0: The population prevalence equals p_0

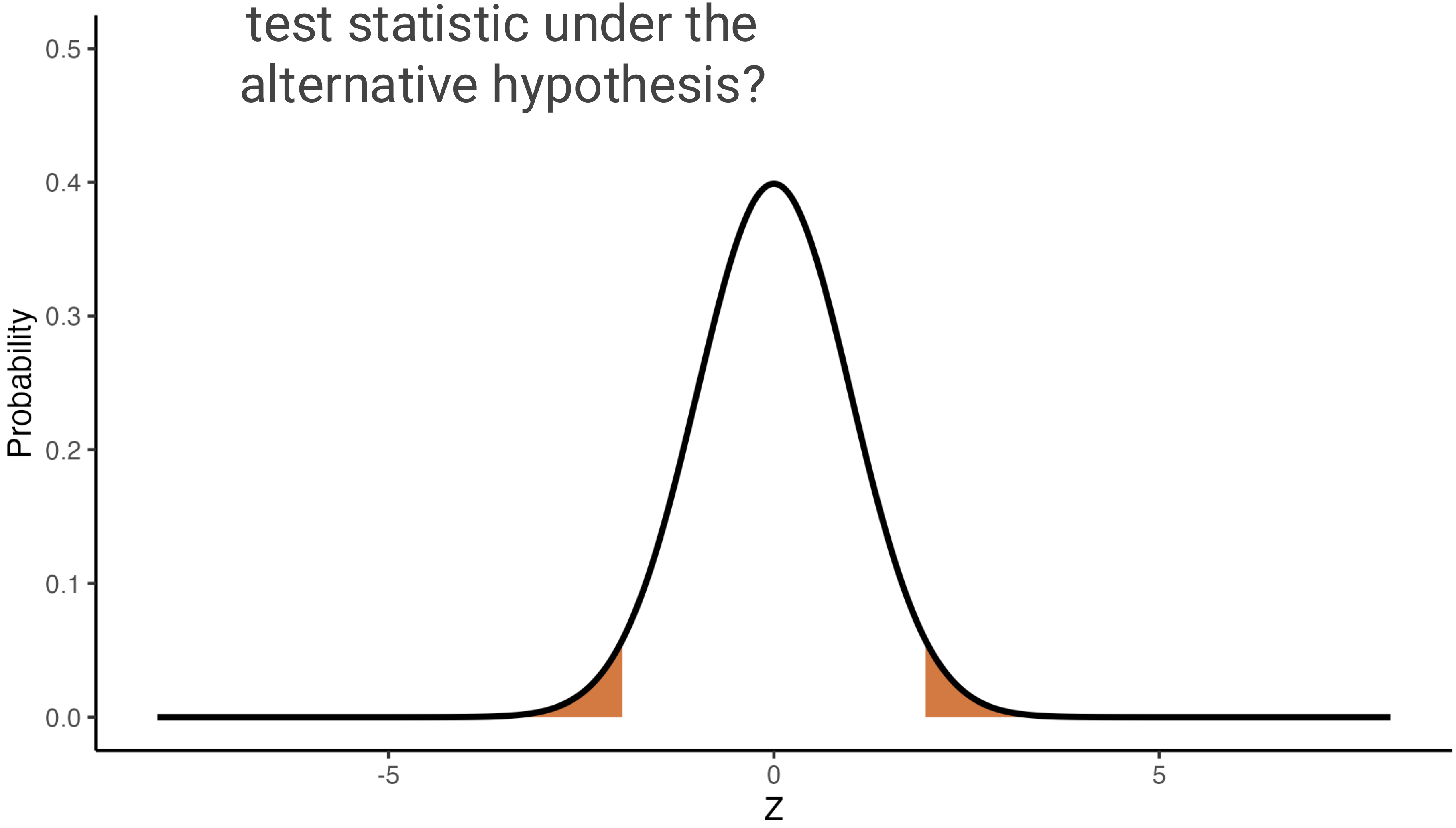
H_1: The population prevalence equals p, which is different from p_0

For example…

I want to test if the prevalence of *pfcrt* K76T mutations is significantly different from 10%. When powering this test, I assume the true prevalence of K76T mutations is 15%.

Power

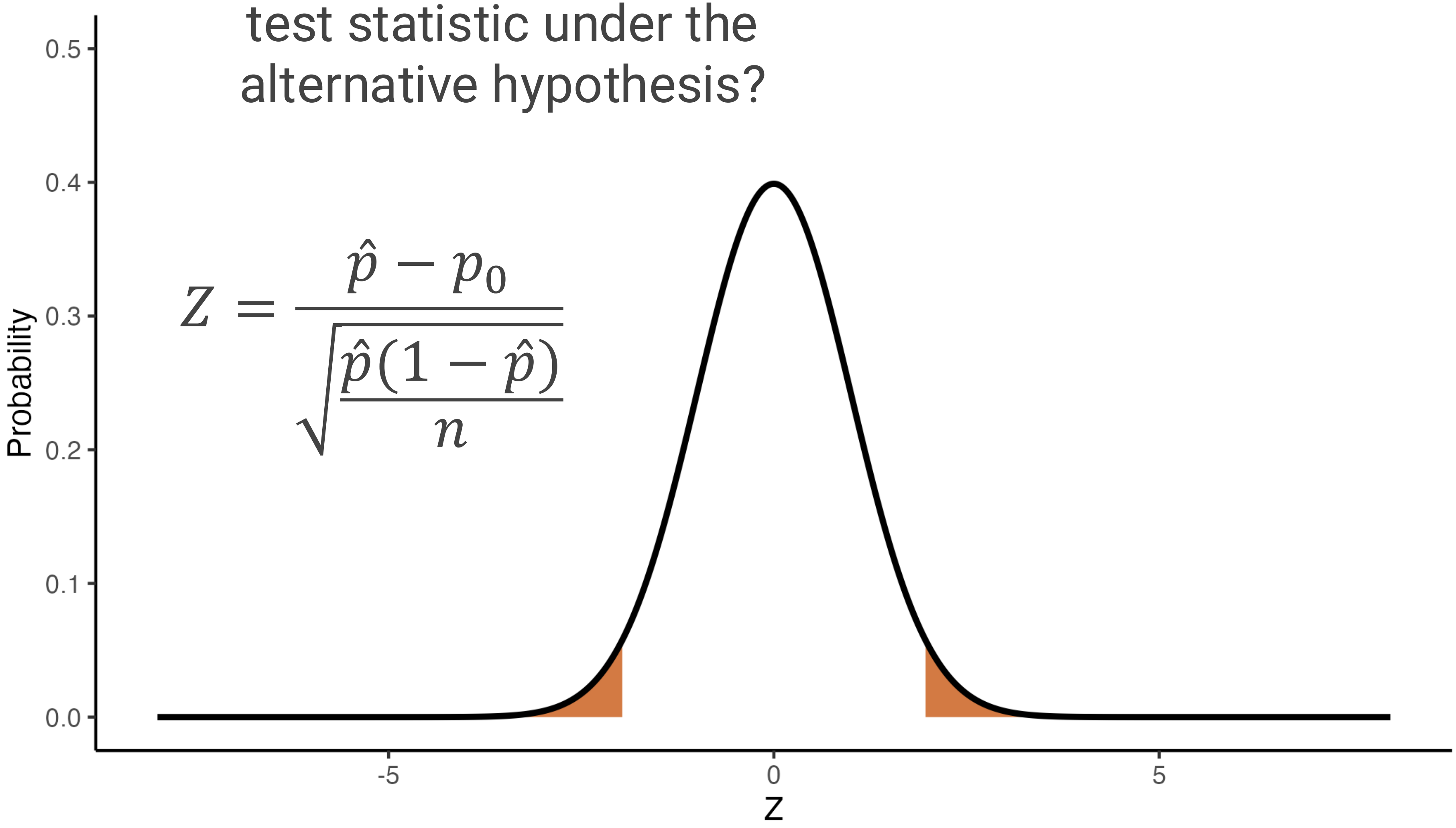
What is the distribution of my test statistic under the alternative hypothesis?



Power

What is the distribution of my test statistic under the alternative hypothesis?

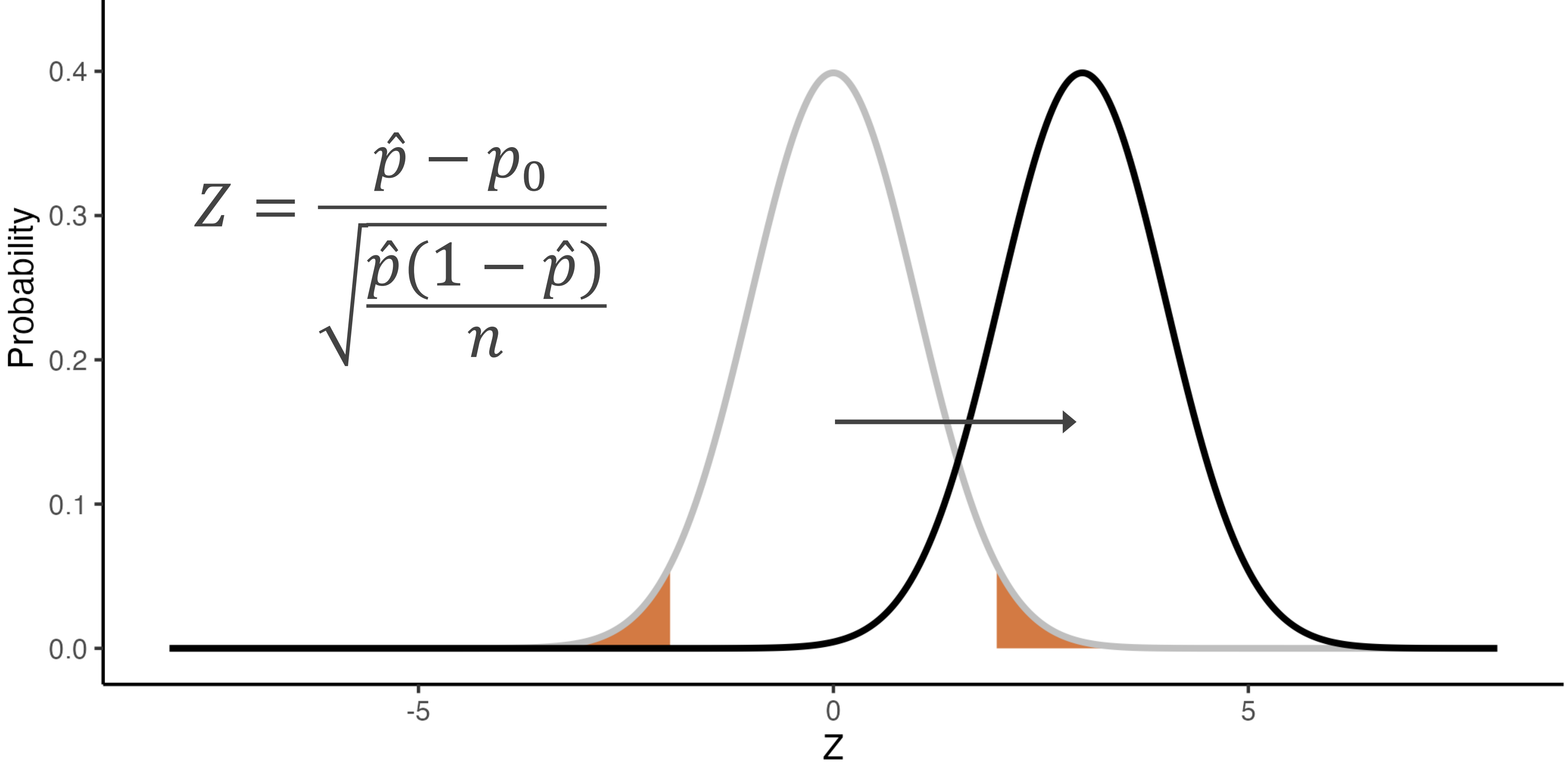
$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$



Power

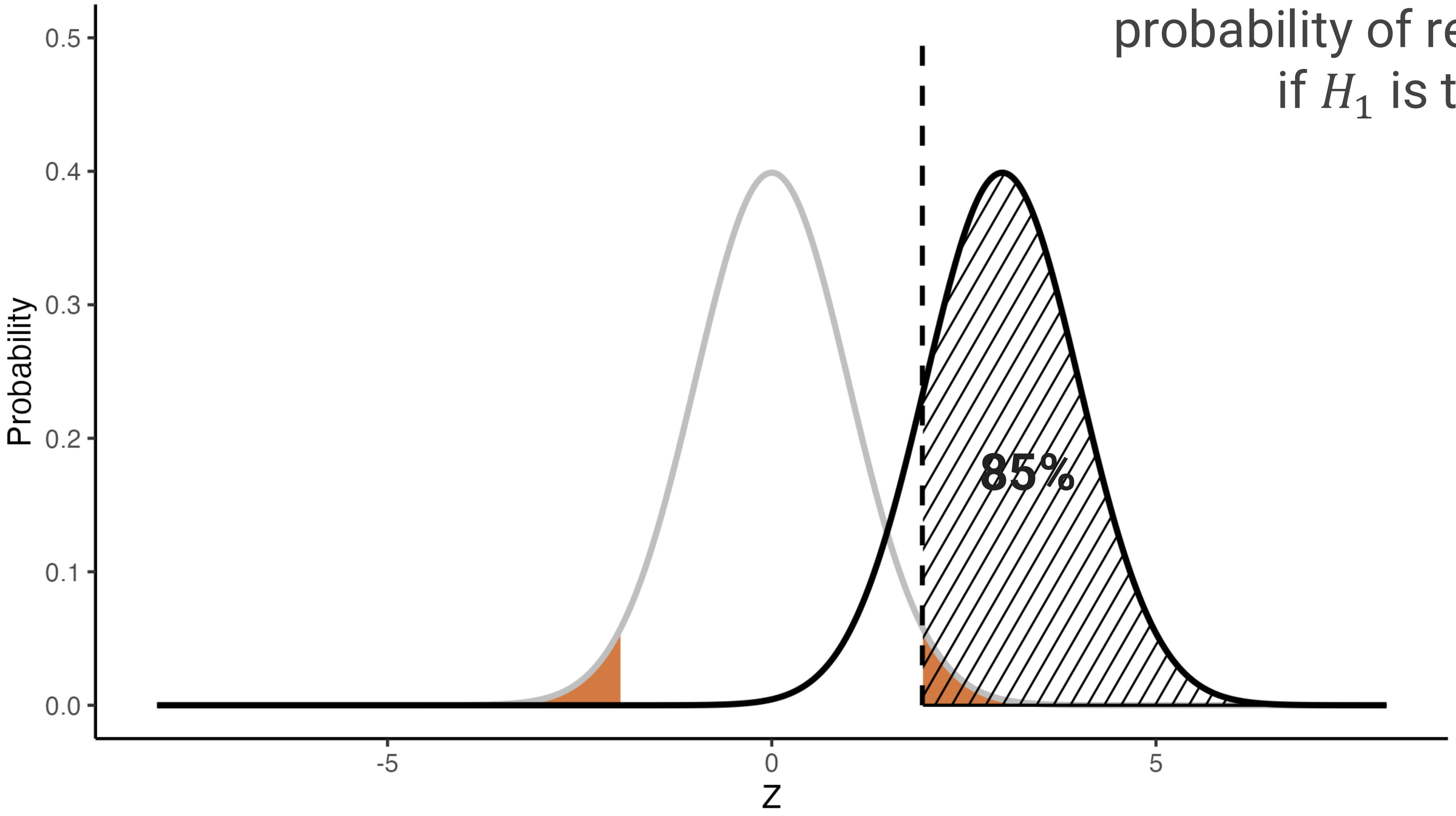
What is the distribution of my test statistic under the alternative hypothesis?

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

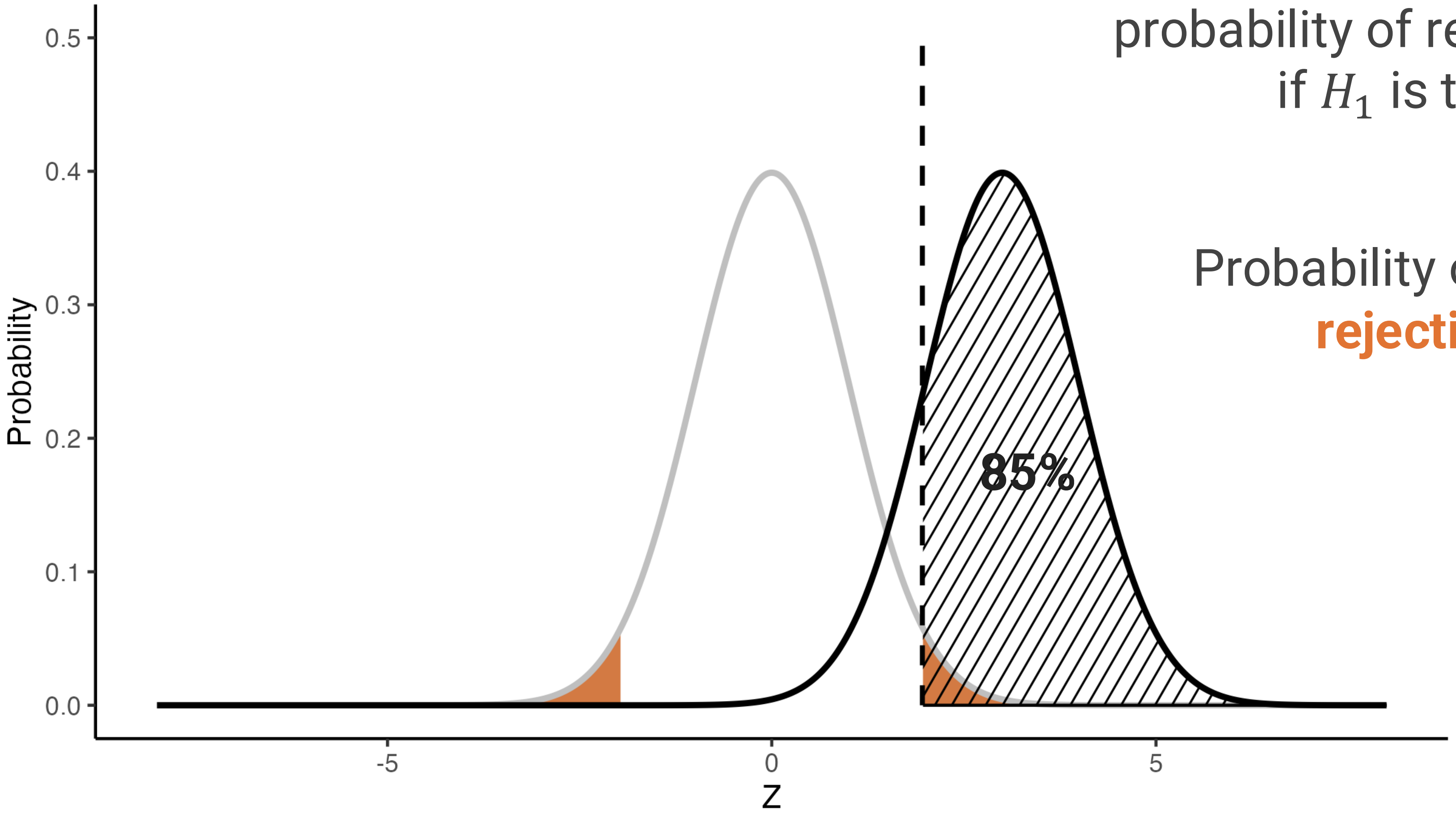


Power

Area under curve gives the probability of rejecting H_0 if H_1 is true



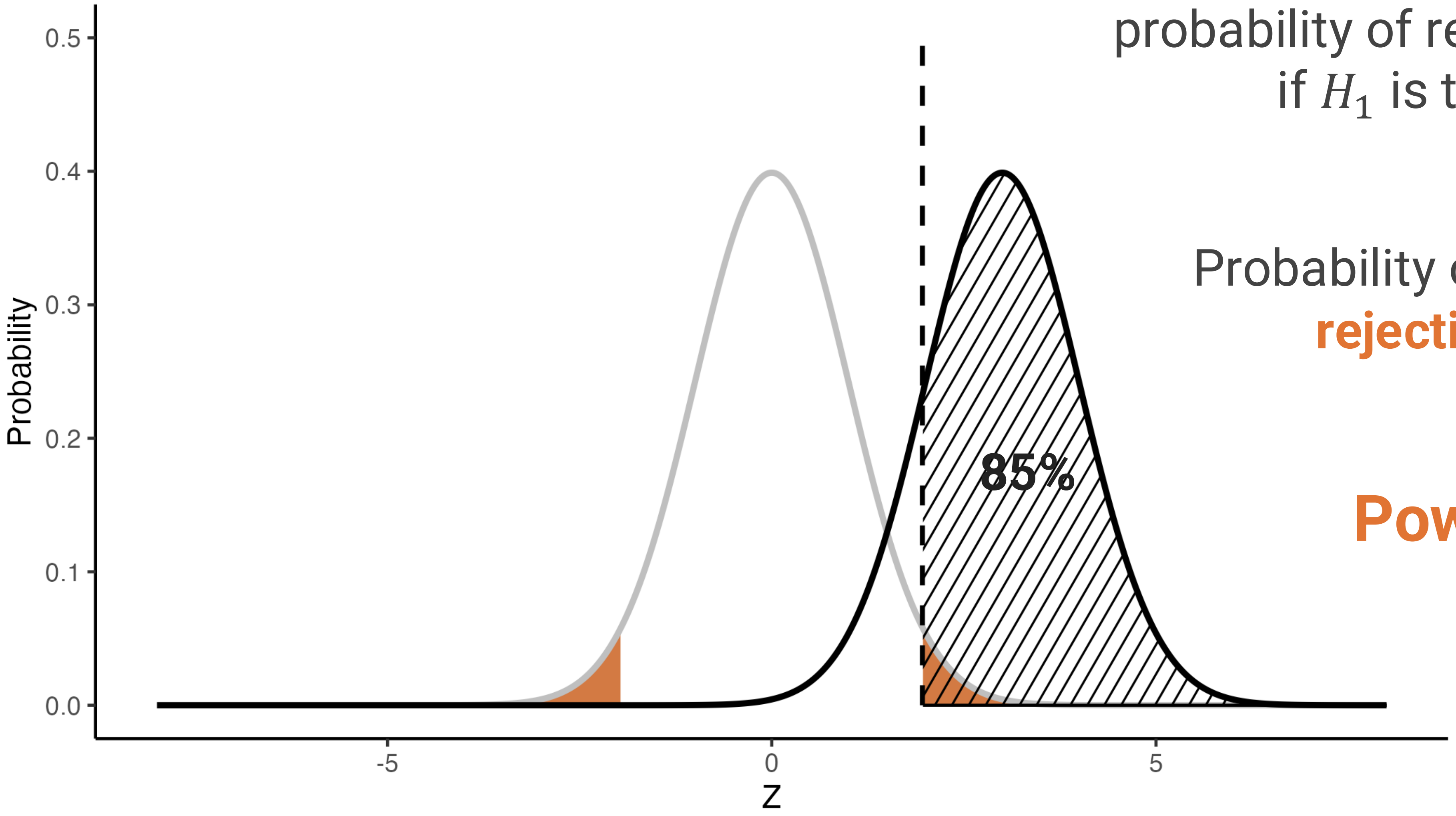
Power



Area under curve gives the probability of rejecting H_0 if H_1 is true

Probability of **correctly rejecting** H_0

Power



Area under curve gives the probability of rejecting H_0 if H_1 is true

Probability of **correctly rejecting** H_0

Power!

Power

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

Power

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

Power

		Conclusion about H_0	
		Fail to reject	Reject
Truth about H_0	True	True negative $1 - \alpha$	False positive α
	False	False negative $1 - \text{Power}$	True positive Power

Power

		Conclusion about H_0	
		Fail to reject	**Reject**
Truth about H_0	**True**	True negative $1-\alpha$	False positive α
	False	False negative $1-\text{Power}$	True positive Power

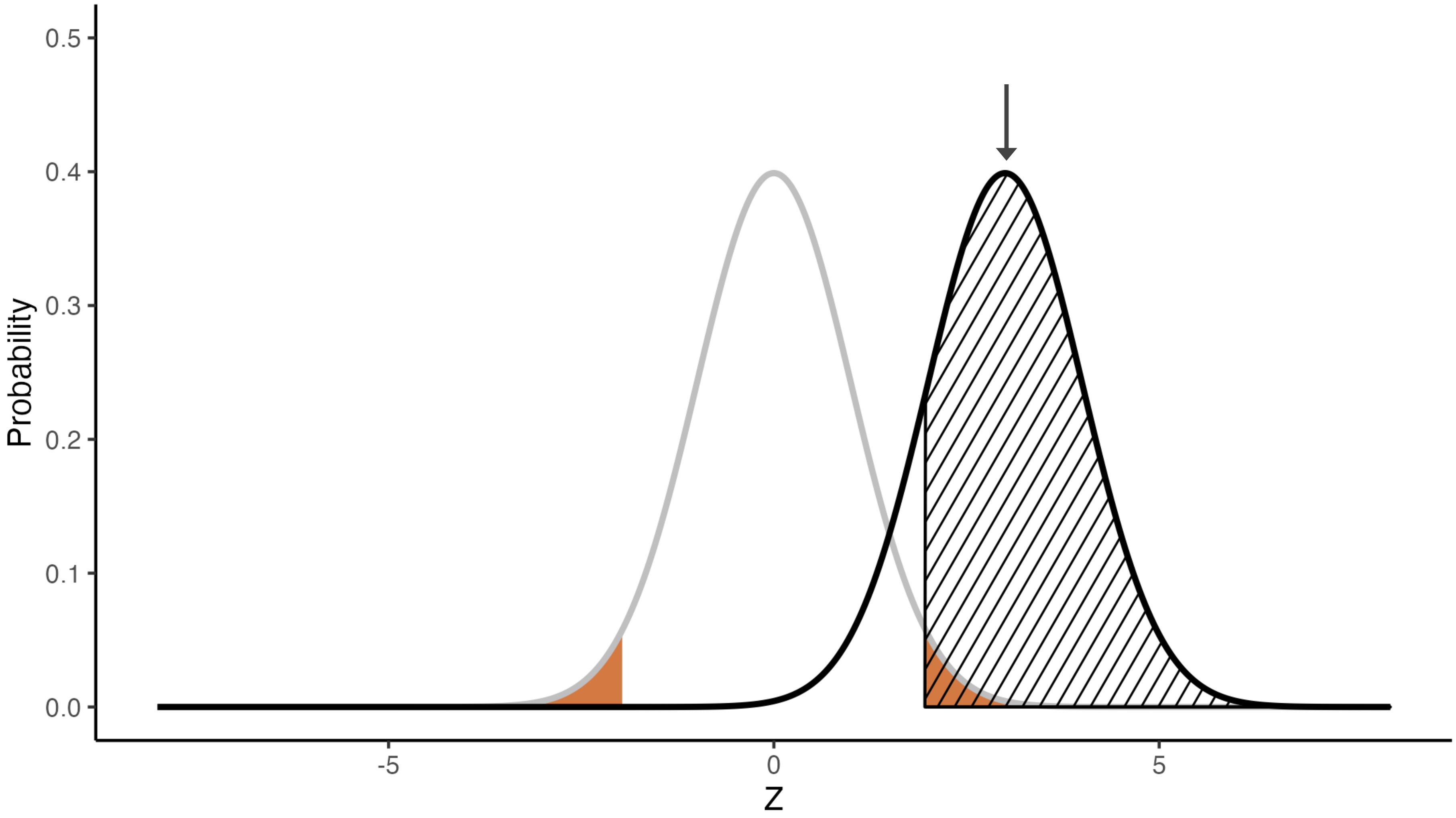
Power is the probability of **correctly rejecting** the null hypothesis. It is the chance that we find something interesting, given that it is there.

Power

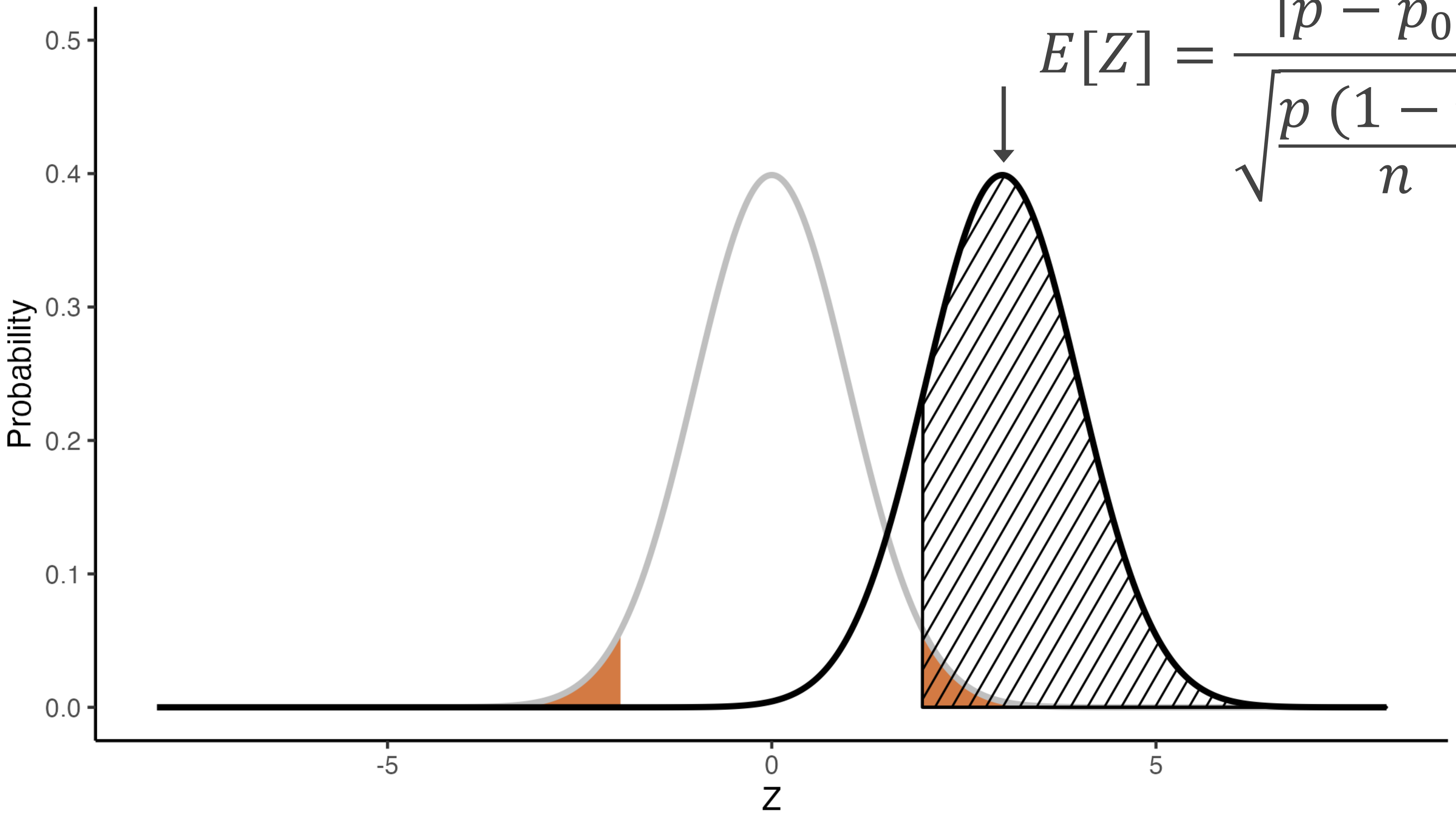
		Conclusion about H_0	
		Fail to reject	Reject
Truth about H_0	True	True negative $1 - \alpha$	False positive α
	False	False negative β	True positive $1 - \beta$

Power is the probability of **correctly rejecting** the null hypothesis. It is the chance that we find something interesting, given that it is there.

How do we calculate power?

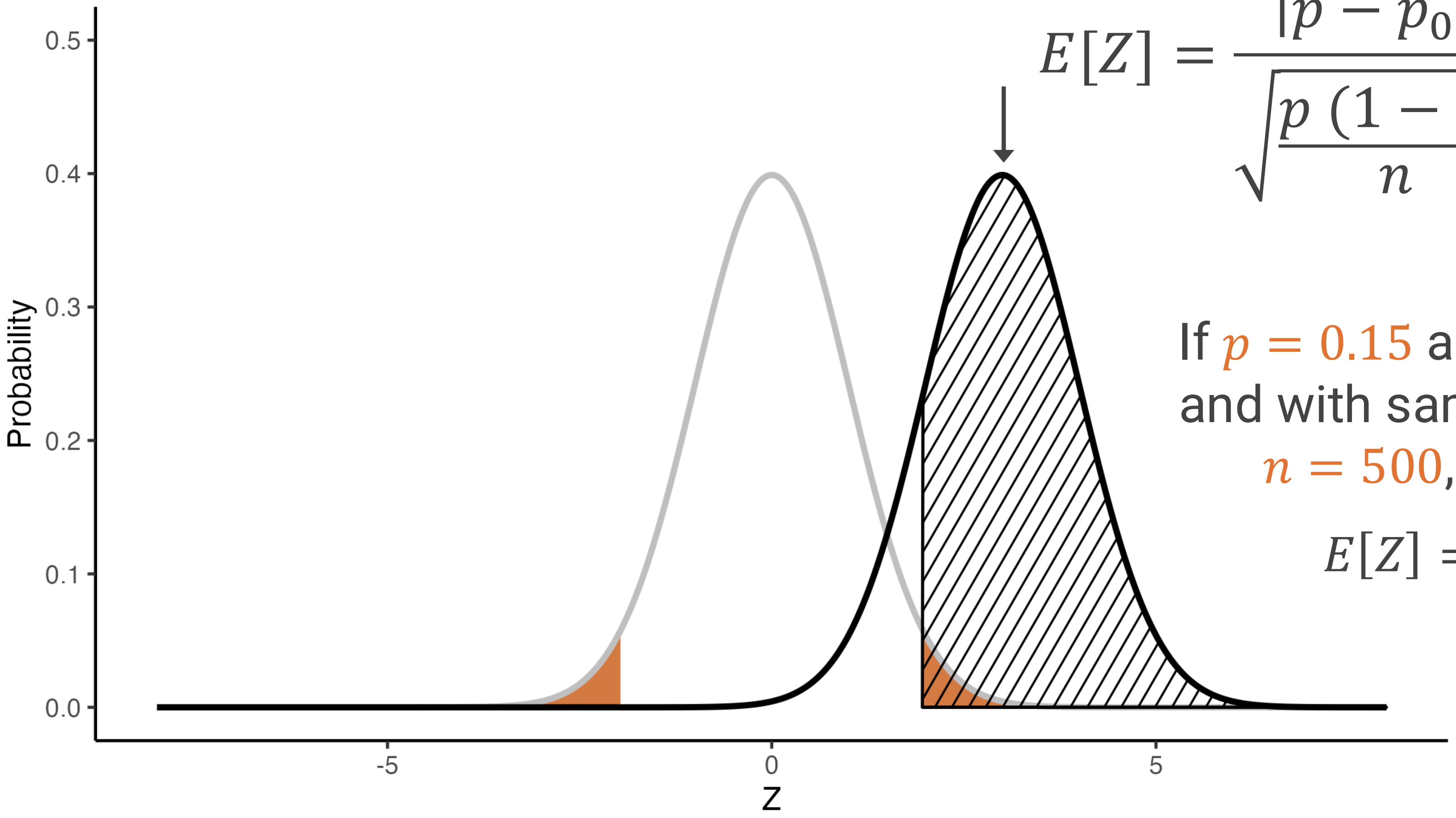


How do we calculate power?



$$E[Z] = \frac{|p - p_0|}{\sqrt{\frac{p(1-p)}{n}}}$$

How do we calculate power?



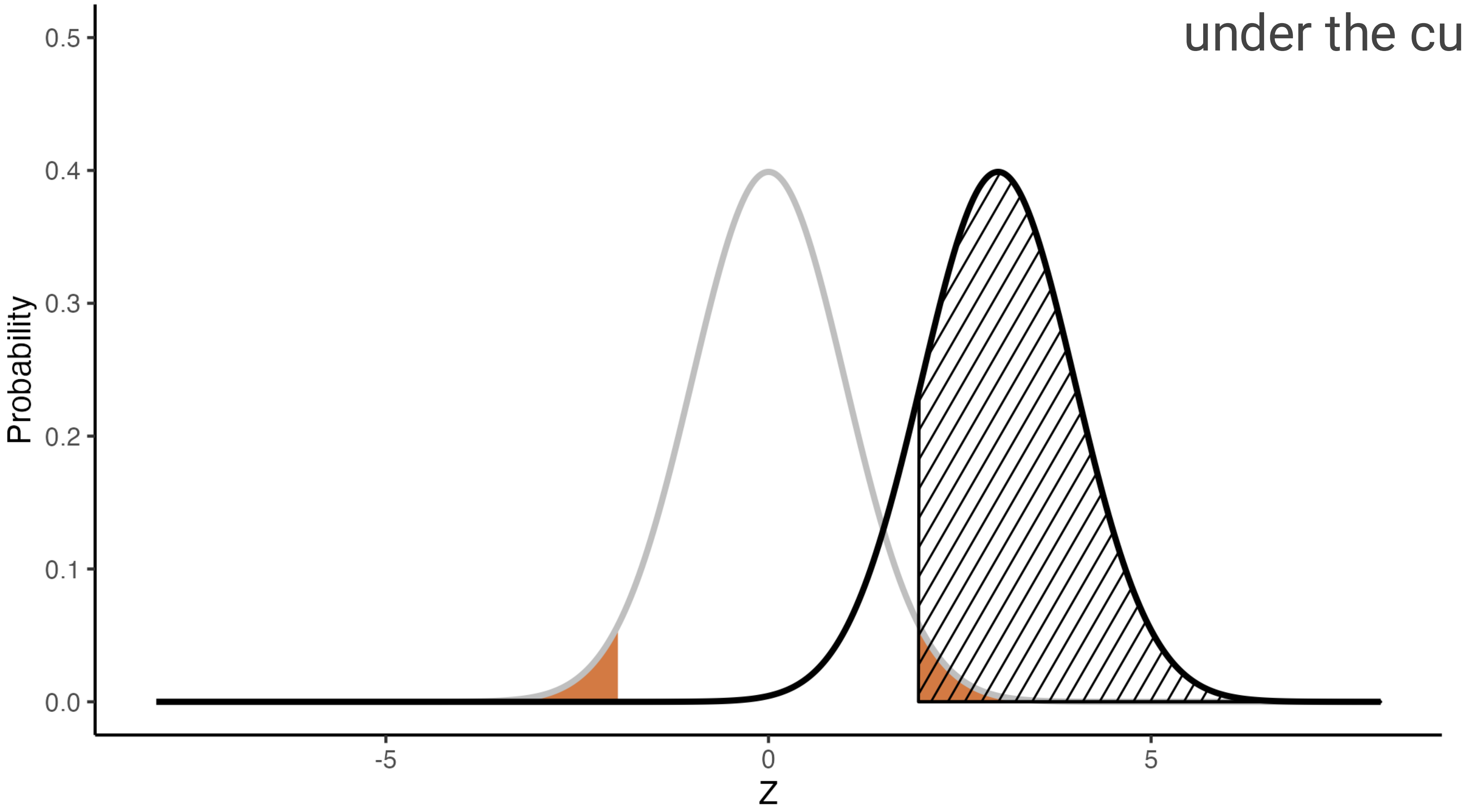
$$E[Z] = \frac{|p - p_0|}{\sqrt{\frac{p\,(1-p)}{n}}}$$

If $p = 0.15$ and $p_0 = 0.1$, and with sample size of $n = 500$, we find

$$E[Z] = 3.13$$

How do we calculate power?

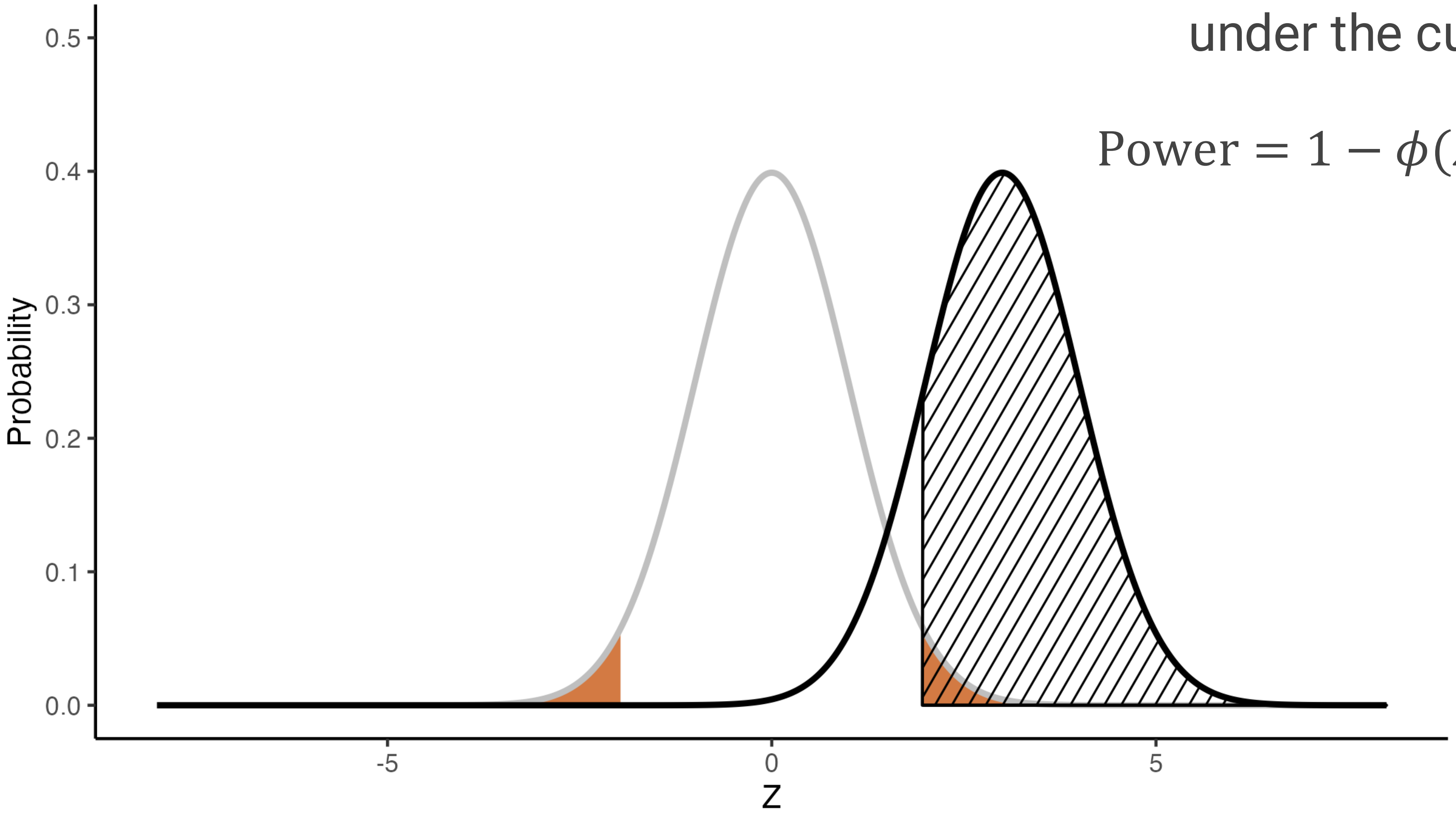
The function $\phi()$ gives us the area under the curve



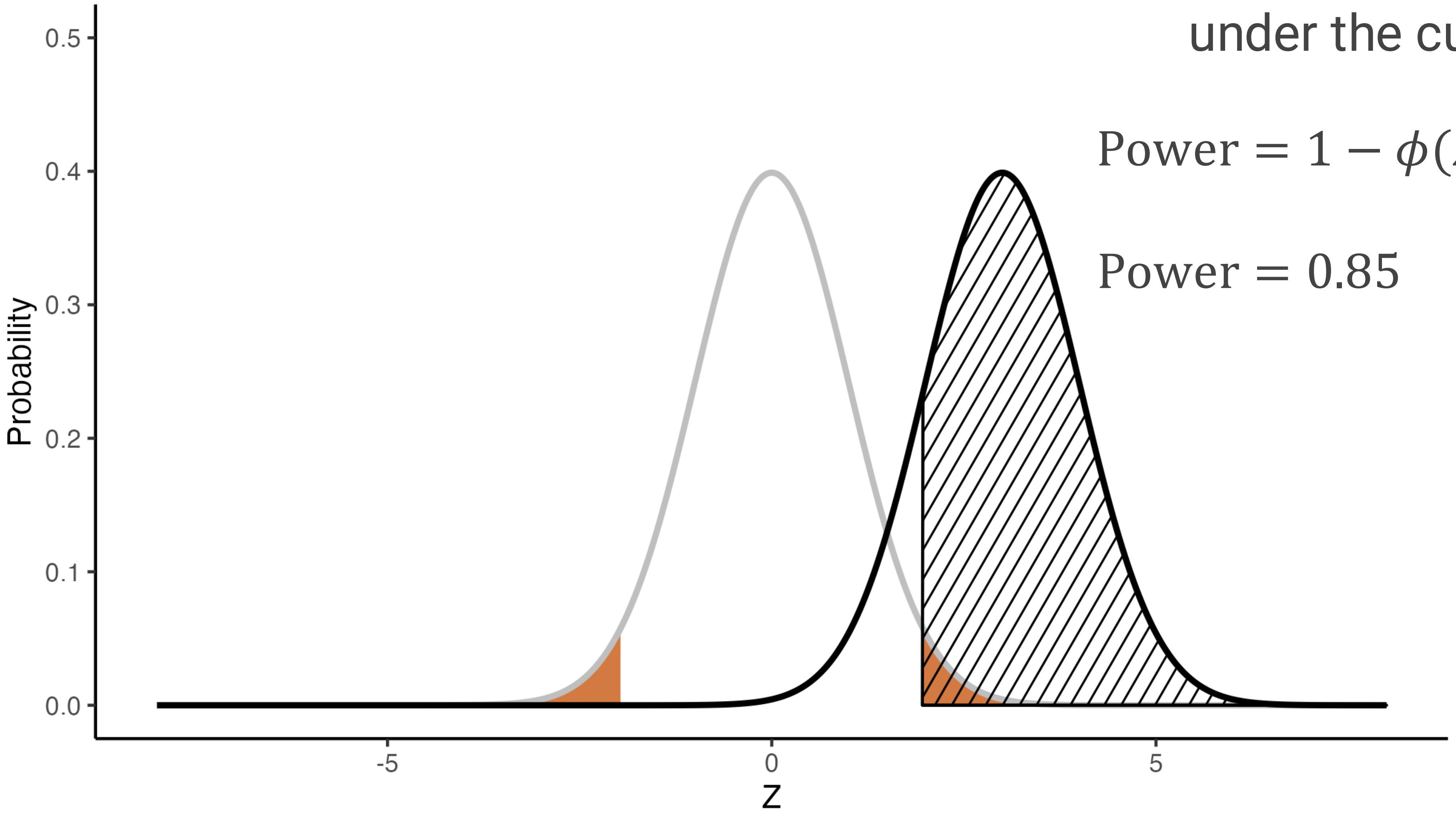
How do we calculate power?

The function $\phi()$ gives us the area under the curve

$$\text{Power} = 1 - \phi(z_{1-\alpha/2} - E[Z])$$



How do we calculate power?



The function $\phi()$ gives us the area under the curve

$$\text{Power} = 1 - \phi(z_{1-\alpha/2} - E[Z])$$

$$\text{Power} = 0.85$$

Power as a function of sample size

$$\text{Power} = 1 - \phi(z_{1-\alpha/2} - E[Z])$$

Power as a function of sample size

$$\text{Power} = 1 - \phi\left(z_{1-\alpha/2} - \frac{|p - p_0|}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Power as a function of sample size

$$\text{Power} = 1 - \phi\left(z_{1-\alpha/2} - \frac{|p - p_0|}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Power varies as a function of sample size

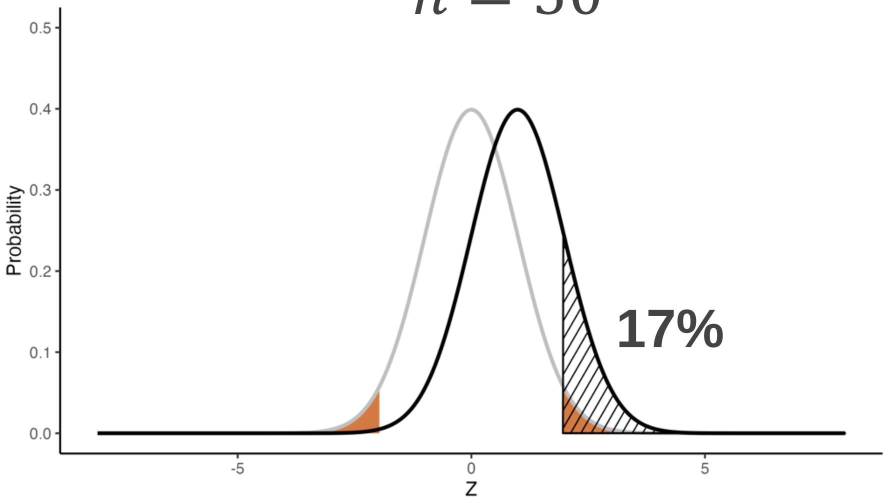
Power as a function of sample size

$$\text{Power} = 1 - \phi\left(z_{1-\alpha/2} - \frac{|p - p_0|}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Power varies as a function of sample size

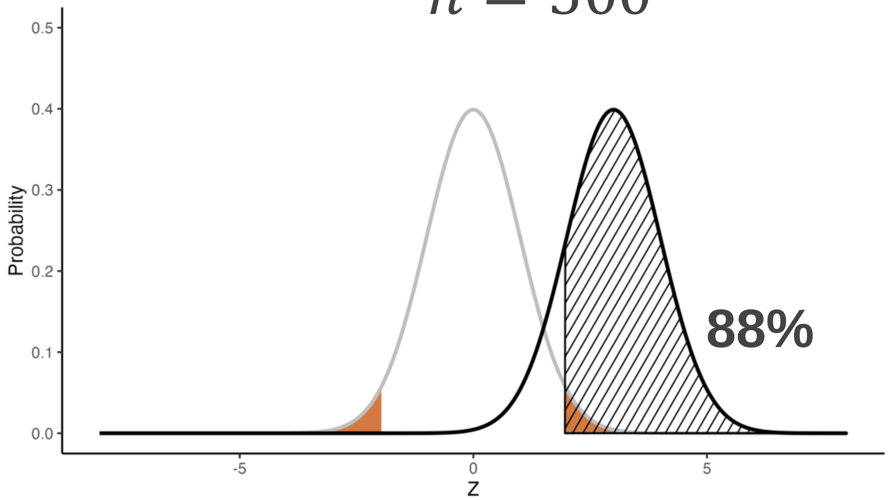
Small sample

$n = 50$

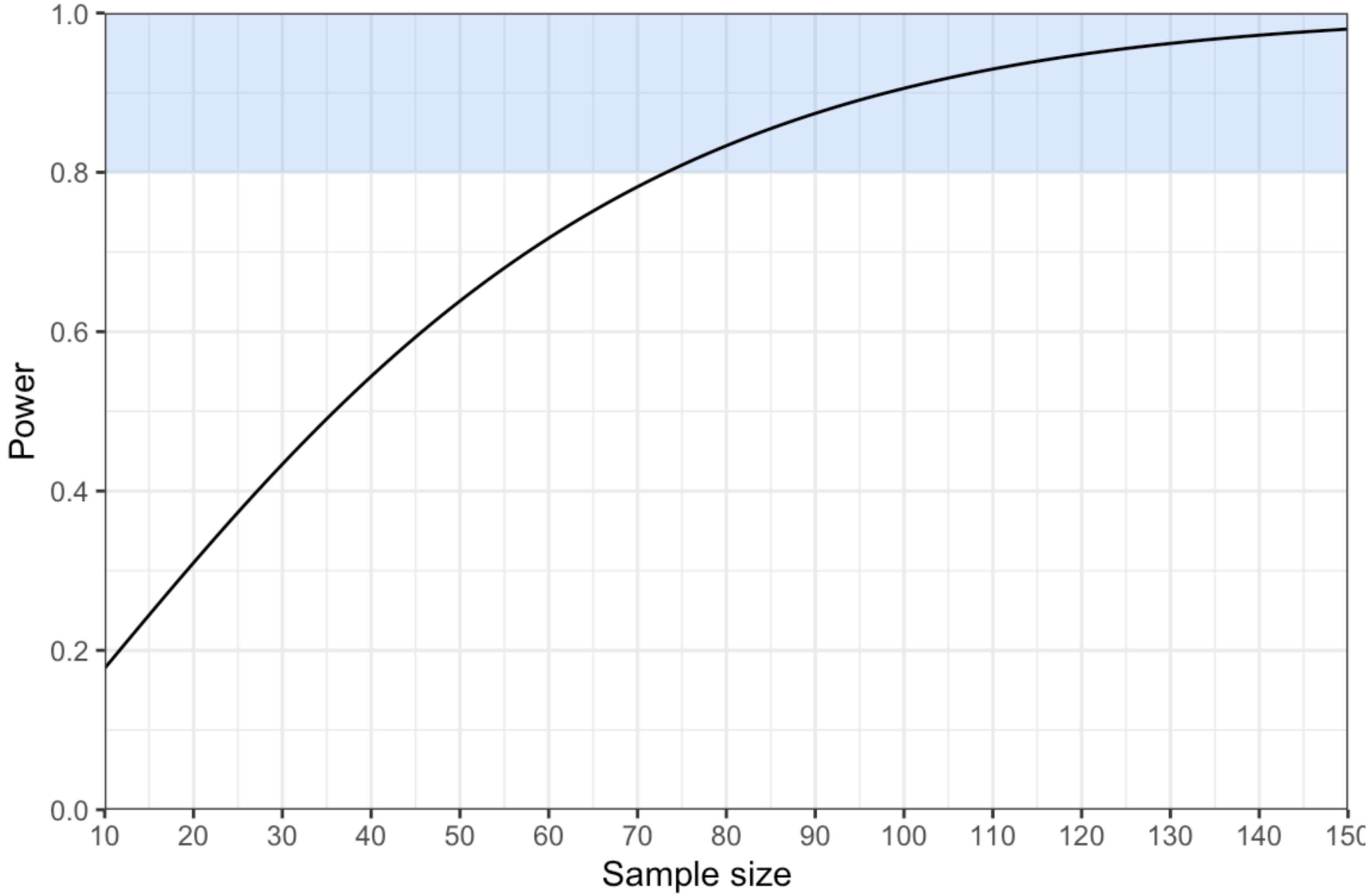


Large sample

$n = 500$

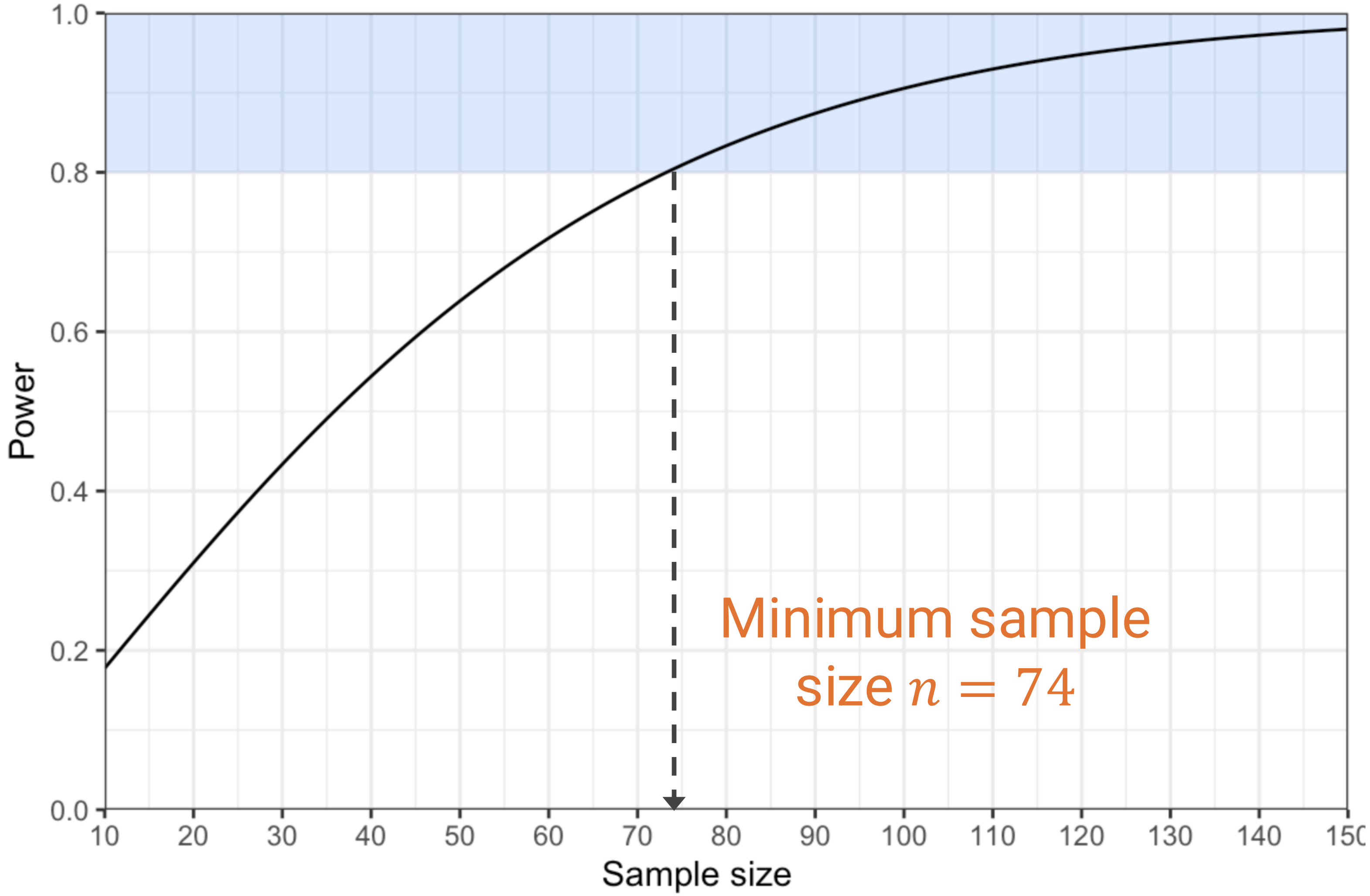


Power curves



80% power

Power curves



80% power

Minimum sample size $n = 74$

Sample size formulae

$$\text{Power} = 1 - \phi\left(z_{1-\alpha/2} - \frac{|p - p_0|}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Can we reverse-engineer this to find the value of n that achieves a target power?

Sample size formulae

$$\text{Power} = 1 - \phi\left(z_{1-\alpha/2} - \frac{|p - p_0|}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Can we reverse-engineer this to find the value of n that achieves a target power?

$$n = \left(z_{1-\beta} + z_{1-\frac{\alpha}{2}}\right)^2 \frac{p(1-p)}{(p - p_0)^2}$$

For 80% power, we find $z_{1-\beta} = 0.84$

Summary

- We can ask questions using **null hypothesis tests**
- A null hypothesis is a statement of **no effect/difference** between groups
- The significance level α controls the **false-positive rate**
- **Power** is the true positive rate. It is the chance of **correctly rejecting the null hypothesis**.
- Power increases with **sample size**. We can use power curves or sample size formulae to choose a value of n

Activity

Format: Interactive R code, accessed through the web

- Short quiz on hypothesis testing
- Test for change in prevalence
- Test for detection of rare *pfk13* variant

Module 3 activity

Format: Interactive R code, accessed through the web

- Short quiz on hypothesis testing
- Test for change in prevalence
- Test for detection of rare *pfk13* variant

Workshop materials

https://mrc-ide.github.io/MMS-SD_workshop/