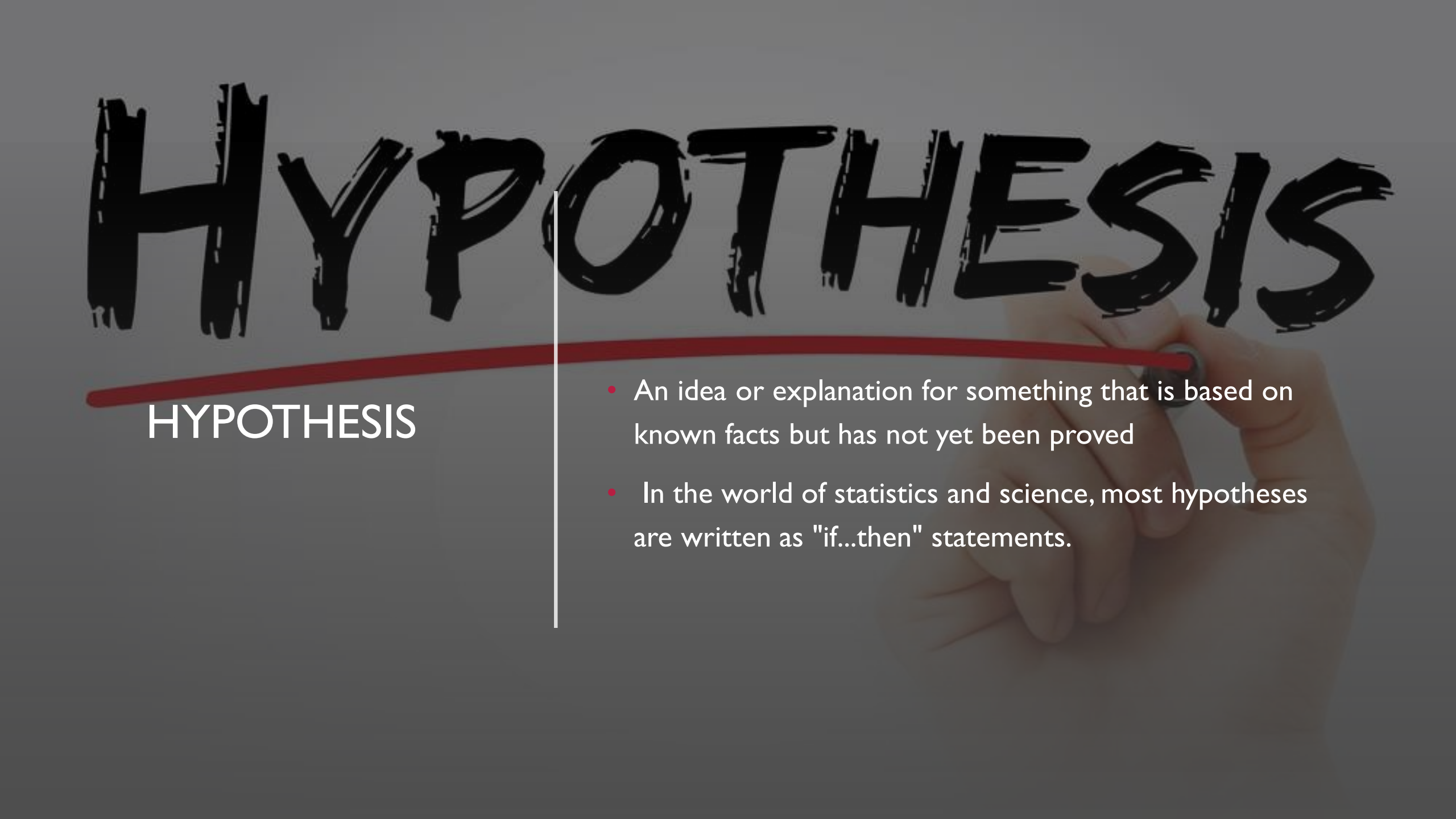
HYPOTHESIS TESTING

LAXMINARAYEN



HYPOTHESIS

- An idea or explanation for something that is based on known facts but has not yet been proved
- In the world of statistics and science, most hypotheses are written as "if...then" statements.

EXAMPLE FOR HYPOTHESIS

- Someone performing experiments on plant growth might report this hypothesis: "If I give a plant an unlimited amount of sunlight, then the plant will grow to its largest possible size."

NULL HYPOTHESIS

- This is the idea that there is no relationship in the population and that the relationship in the sample reflects only sampling error.
- **Also Known As: H0, no-difference hypothesis**

EXAMPLE FOR NULL HYPOTHESIS

- To write a null hypothesis, first start by asking a question. Rephrase that question in a form that assumes no relationship between the variables.

Question	**Null Hypothesis**
Are teens better at math than adults?	Age has no effect on mathematical ability.
Does taking aspirin every day reduce the chance of having a heart attack?	Taking aspirin daily does not affect heart attack risk.

EXAMPLES OF NULL HYPOTHESIS

Do teens use cell phones to access the internet more than adults?

EXAMPLES OF NULL HYPOTHESIS

Do teens use cell phones to access the internet more than adults?

- Age has no effect on how cell phones are used for internet access.

EXAMPLES OF NULL HYPOTHESIS

Do teens use cell phones to access the internet more than adults?

- Age has no effect on how cell phones are used for internet access.

Do cats care about the color of their food?

EXAMPLES OF NULL HYPOTHESIS

Do teens use cell phones to access the internet more than adults?

- Age has no effect on how cell phones are used for internet access.

Do cats care about the color of their food?

- Cats express no food preference based on color.

EXAMPLES OF NULL HYPOTHESIS

Do teens use cell phones to access the internet more than adults?

- Age has no effect on how cell phones are used for internet access.

Do cats care about the color of their food?

- Cats express no food preference based on color.

Does chewing willow bark relieve pain?

EXAMPLES OF NULL HYPOTHESIS

Do teens use cell phones to access the internet more than adults?

- Age has no effect on how cell phones are used for internet access.

Do cats care about the color of their food?

- Cats express no food preference based on color.

Does chewing willow bark relieve pain?

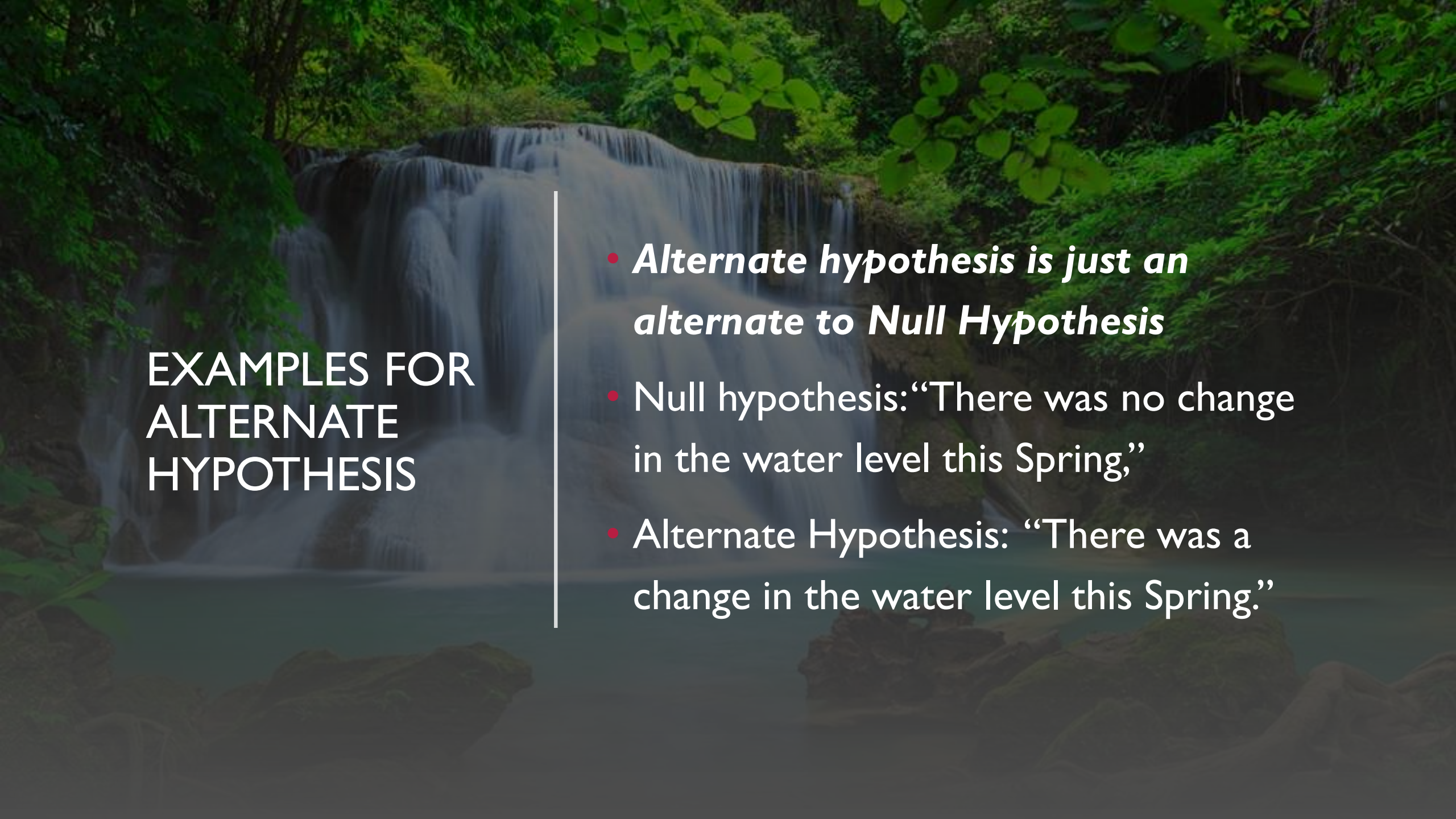
- There is no difference in pain relief after chewing willow bark versus taking a placebo.

NULL HYPOTHESIS

- Informally, the null hypothesis is that the sample relationship "occurred by chance."
- The pattern we find the sample does not occur in the population

ALTERNATE HYPOTHESIS

- This is the idea that there is a relationship in the population and that the relationship in the sample reflects this relationship in the population.
- The **alternative hypothesis** (often symbolized as H_1).



EXAMPLES FOR ALTERNATE HYPOTHESIS

- ***Alternate hypothesis is just an alternate to Null Hypothesis***
- Null hypothesis: "There was no change in the water level this Spring,"
- Alternate Hypothesis: "There was a change in the water level this Spring."

EVERY STATISTICAL RELATIONSHIP CAN BE PUT IN TWO WAYS

It might have occurred by chance – H0

Or it might reflect a relationship in the population. – H1

WHY TEST THE NULL HYPOTHESIS?

You may be wondering why you would want to test a hypothesis just to find it false.

Why not just test an alternate hypothesis and find it true?

WHY TEST THE NULL HYPOTHESIS?

- It turns out it's much easier to disprove a hypothesis than to ever prove one.

FRAMING THE HYPOTHESIS

- At the start of the experiment, the null hypothesis is assumed to be true.
- If the data fails to support the null hypothesis, only then can we look to an alternative hypothesis

EXAMPLE FOR FRAMING THE HYPOTHESIS

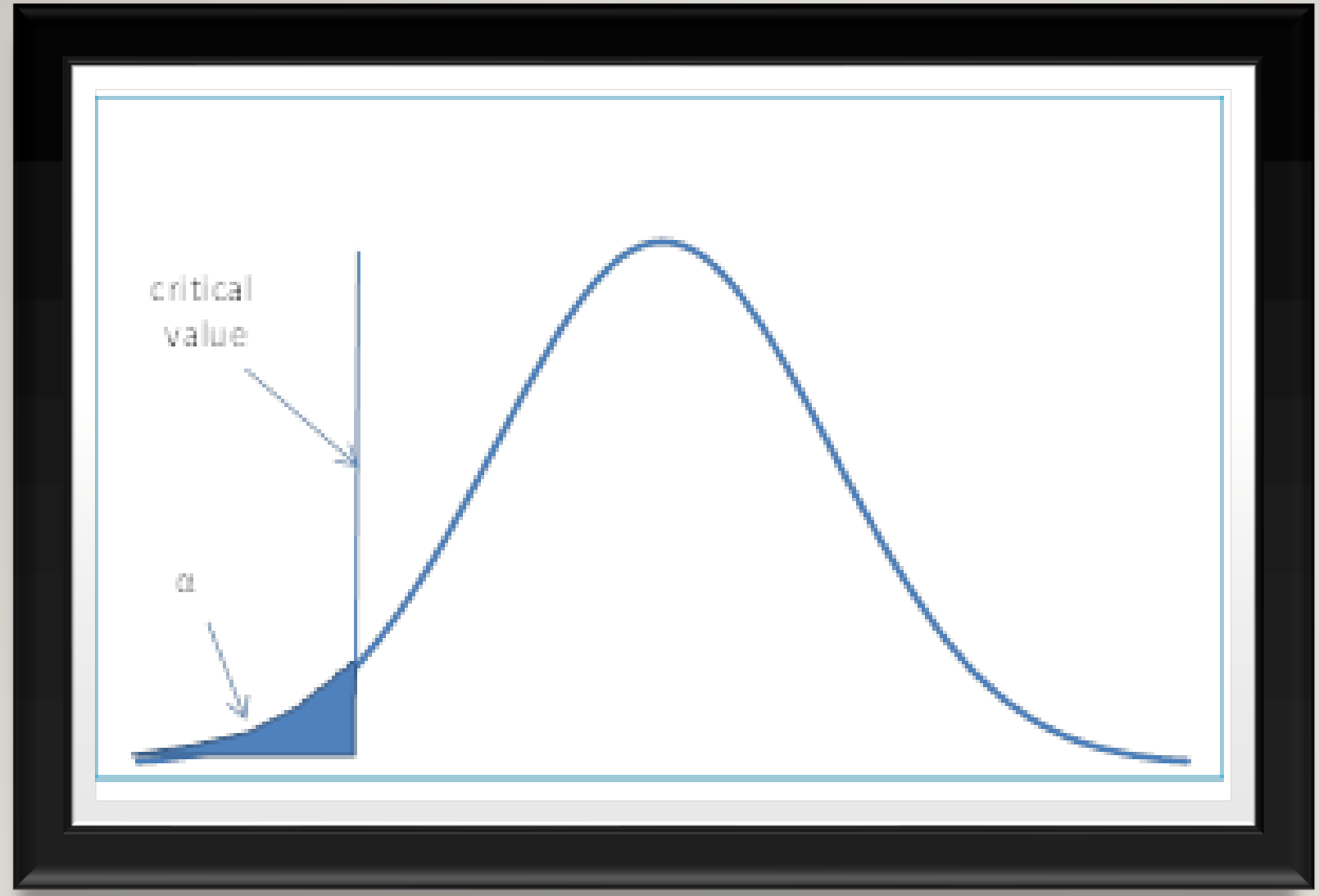
- If testing something assumed to be true,
- the null hypothesis can reflect the assumption:
- Claim: "Our product has an average shipping weight of 3.5kg."
- Null hypothesis: average weight = 3.5kg
- Alternate hypothesis: average weight ≠ 3.5kg

EXAMPLE FOR FRAMING THE HYPOTHESIS

- If testing a claim we want to be true,
- but can't assume, we test its opposite:
- Claim: "This prep course improves test scores."
- Null hypothesis: old scores ≥ new scores
- Alternate hypothesis: old scores < new scores

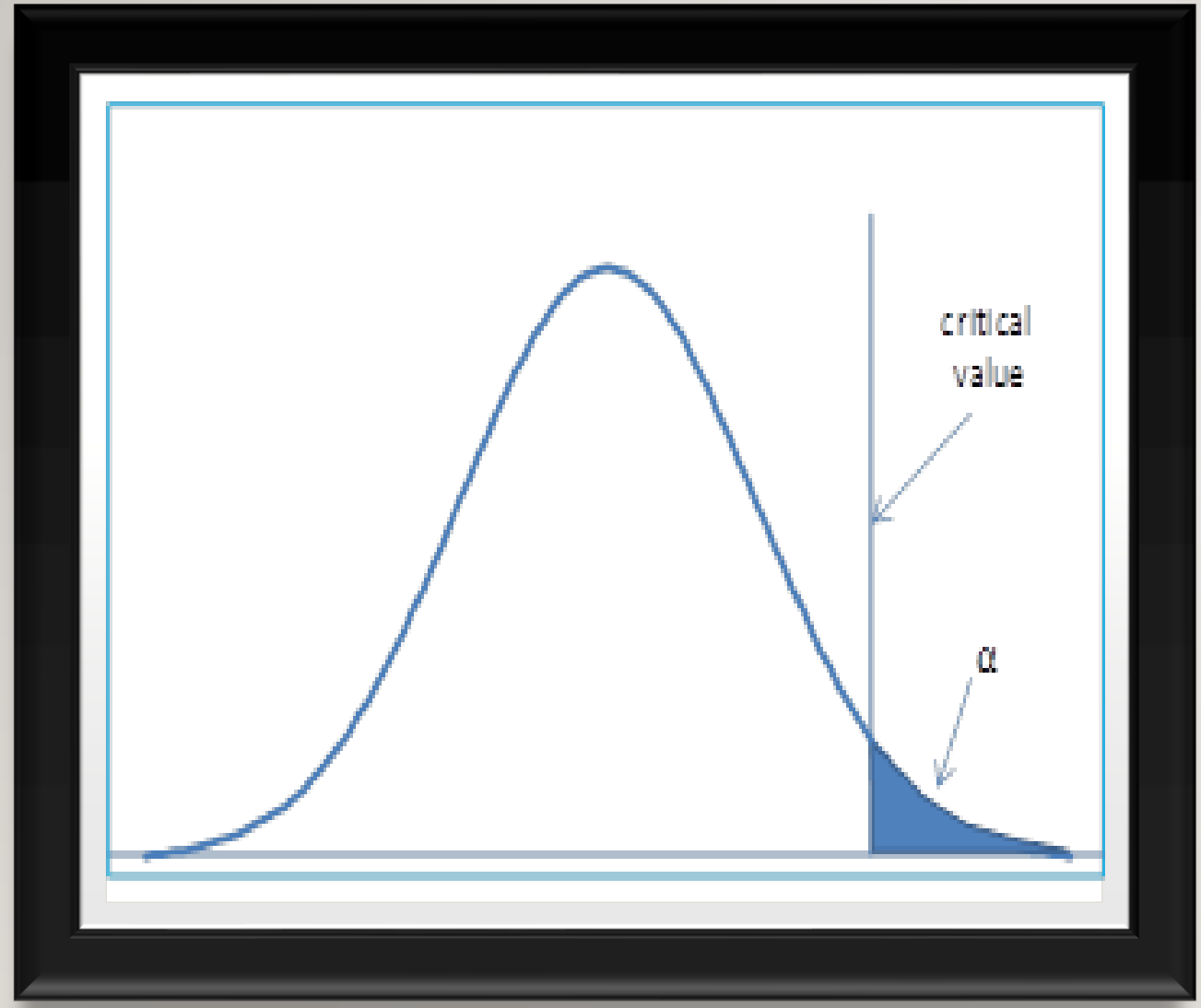
FRAMING THE HYPOTHESIS

- The null hypothesis should contain an equality (=, ≤ ,≥):
- average shipping weight = 3.5kg $H_0: \mu = 3.5$
- The alternate hypothesis should not have an equality (≠,<,>):
- average shipping weight ≠ 3.5kg $H1: \mu \neq 3.5$



LEFT TAIL TEST

LEFT TAIL: WHEN THE HA < (HAS A LESS THAN SYMBOL) WE GO WITH THE LEFT TAIL TEST

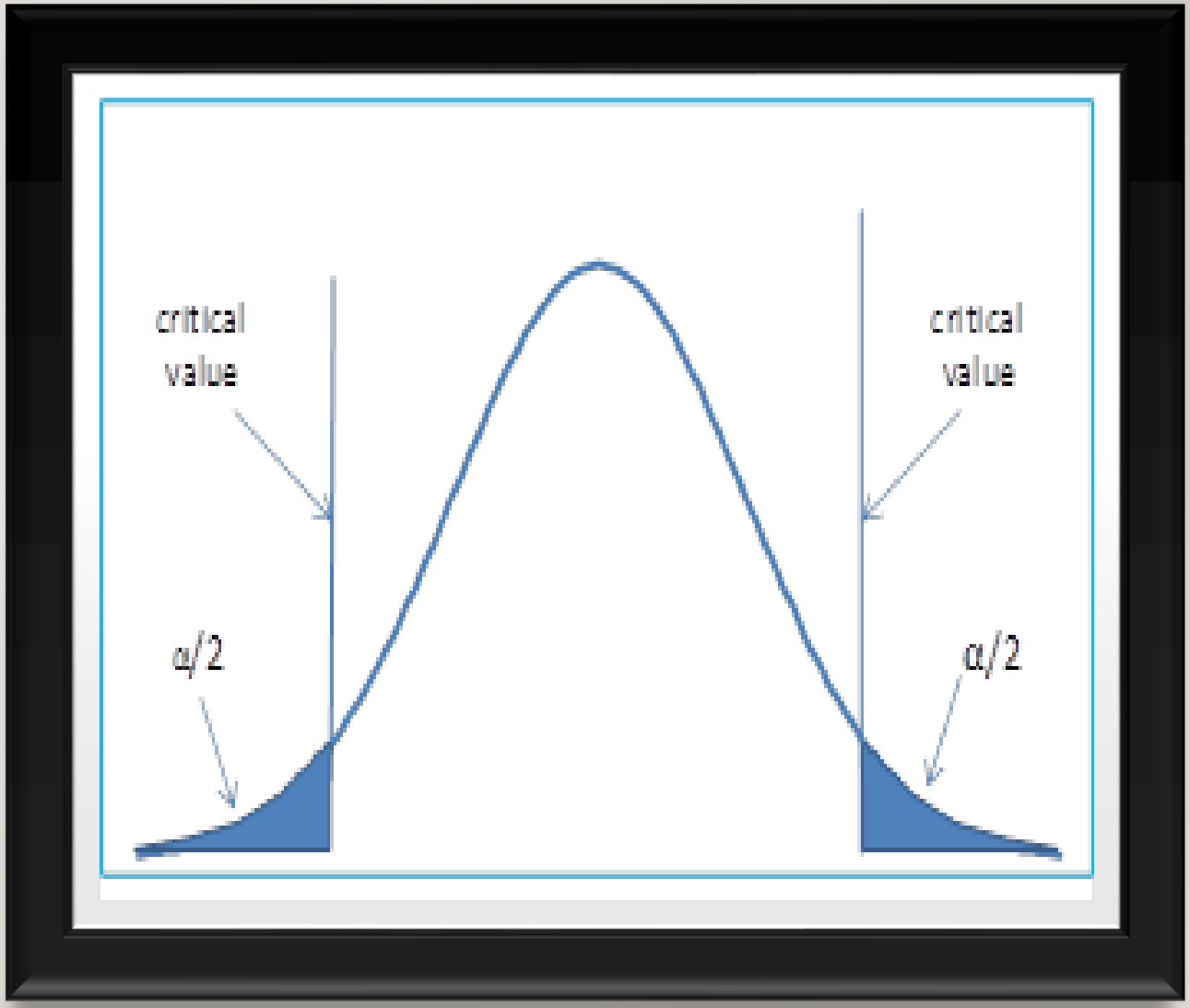


RIGHT TAIL TEST

- LEFT TAIL: WHEN THE HA > (HAS A GREATER THAN SYMBOL) WE GO WITH THE LEFT TAIL TEST

TWO TAIL TEST

LEFT TAIL: WHEN THE HA ≠ (HAS A NOT EQUAL TO SYMBOL) WE GO WITH THE LEFT TAIL TEST



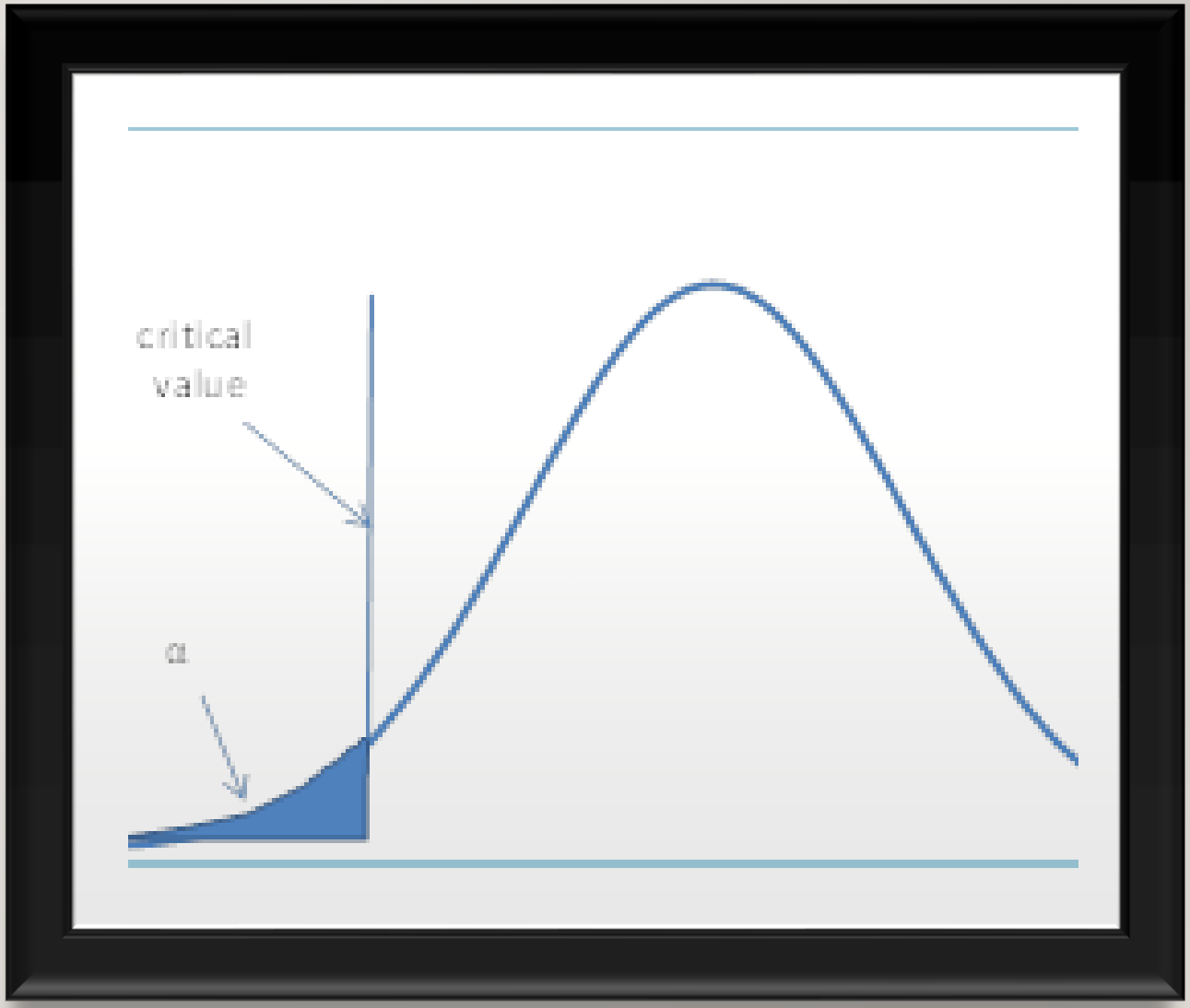
EXAMPLE FOR TAILS

- Consider this Push pins packet
- They Claim to have 100 rose gold pins in this packet

EXAMPLE FOR TAILS

- Let us say we are tricked and we have less than 100 pins in there.

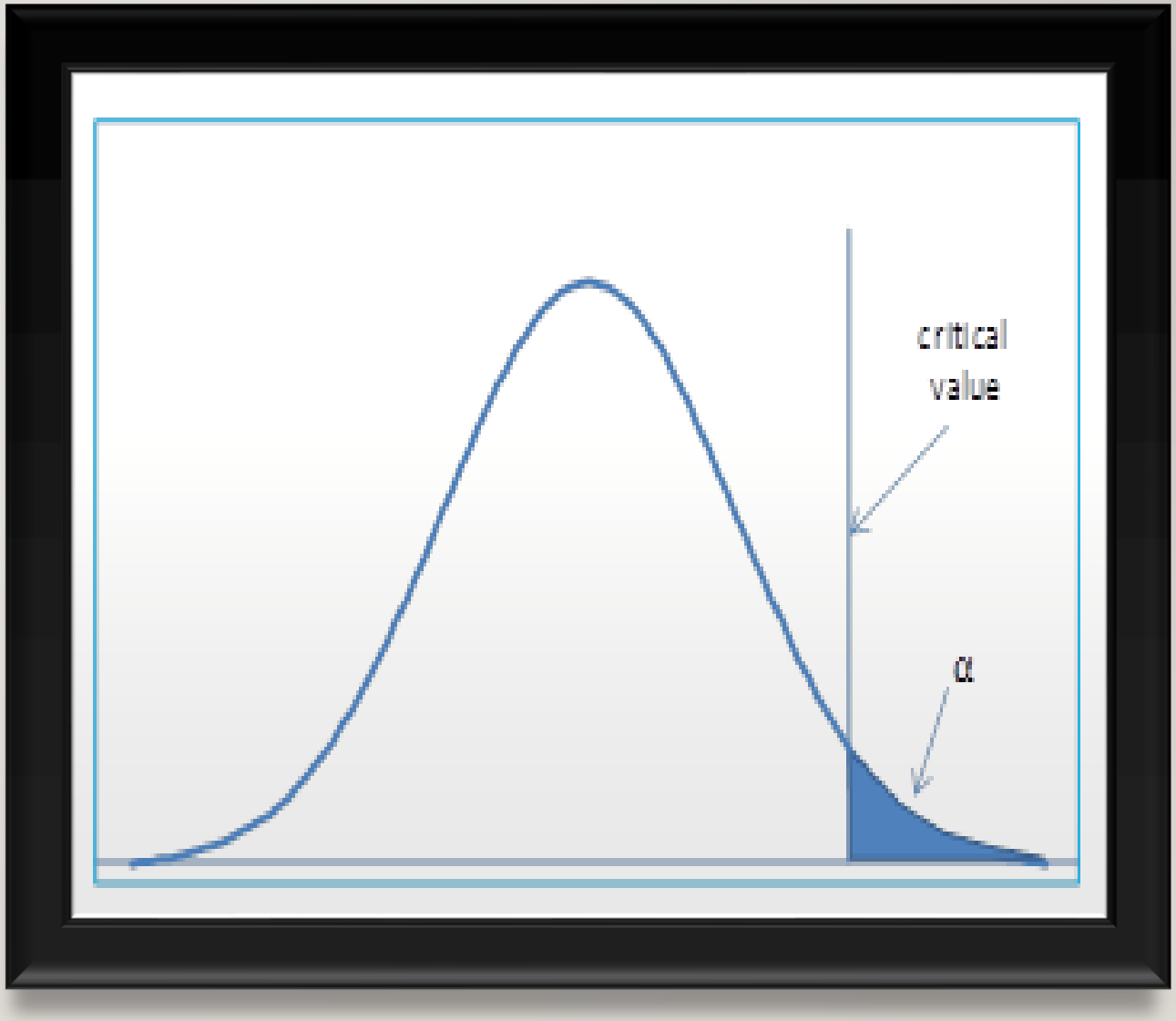
Then it is less than 100 left tail test!!



EXAMPLE FOR TAILS

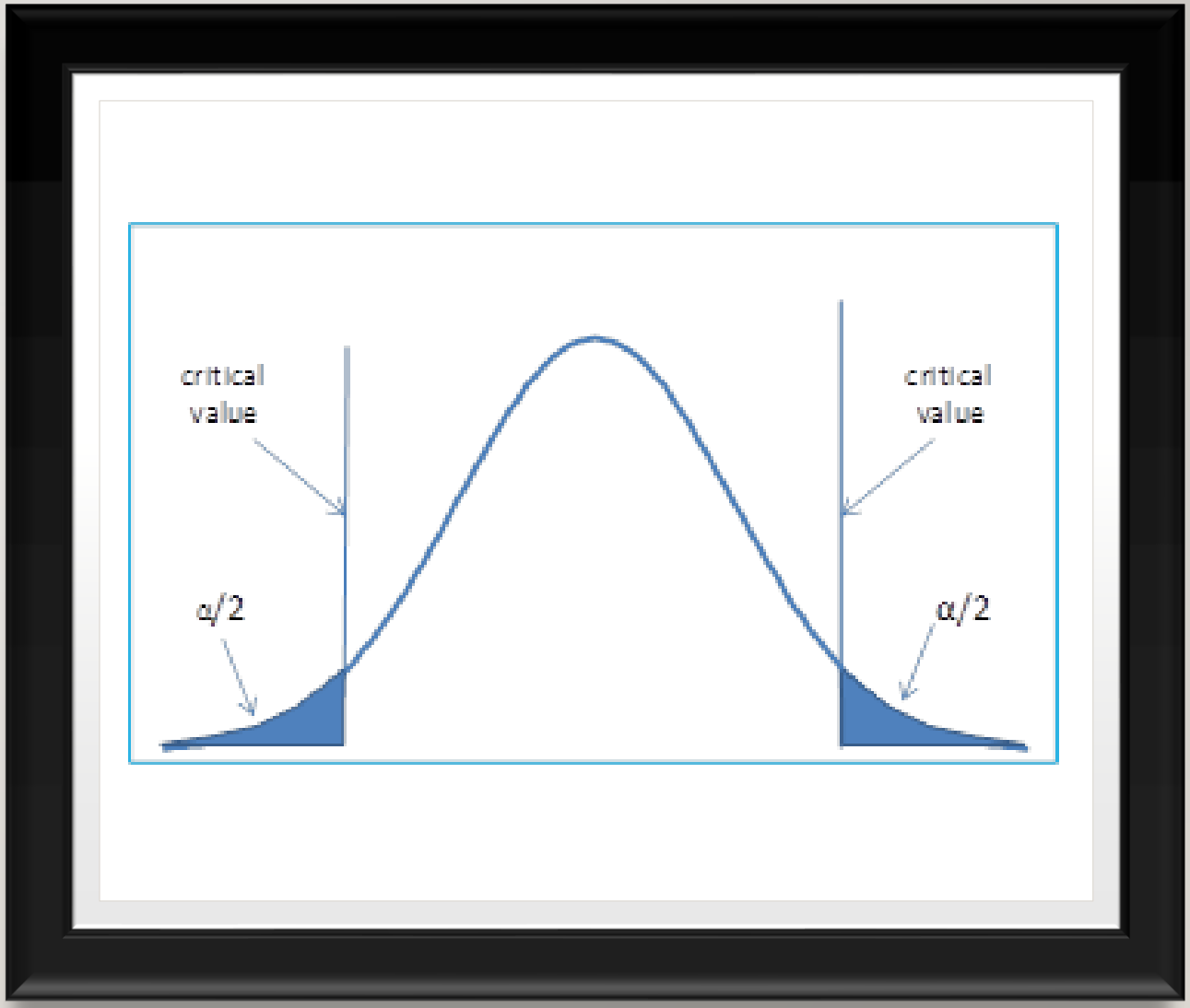
- Let us say we are lucky and we have more than 100 pins in there.

Then it is greater than 100 right tail test!!



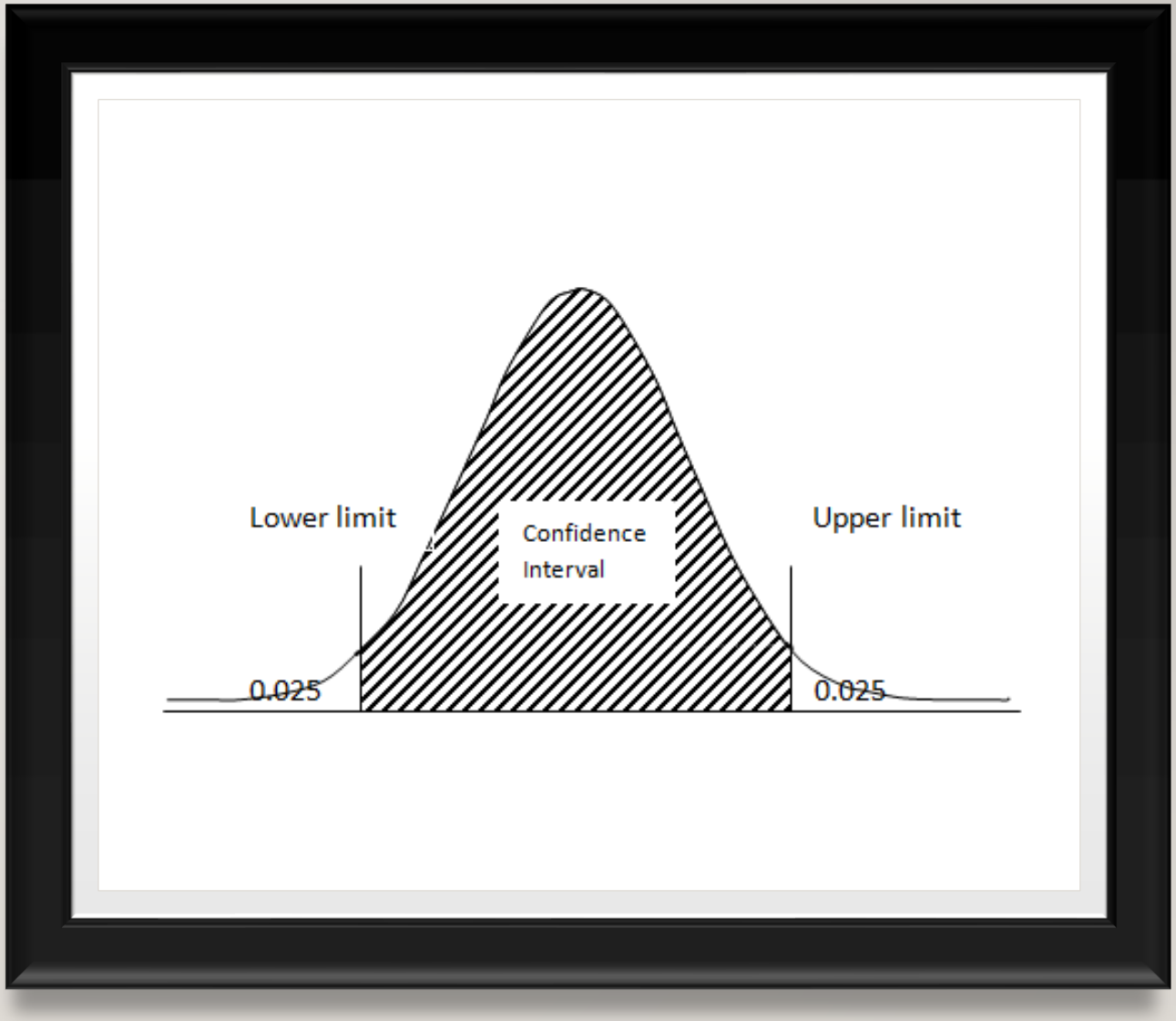
EXAMPLE FOR TAILS

- Let us say there are not 100 pins in it.
- There I do not define if its less than or greater than. Then it is two tail.

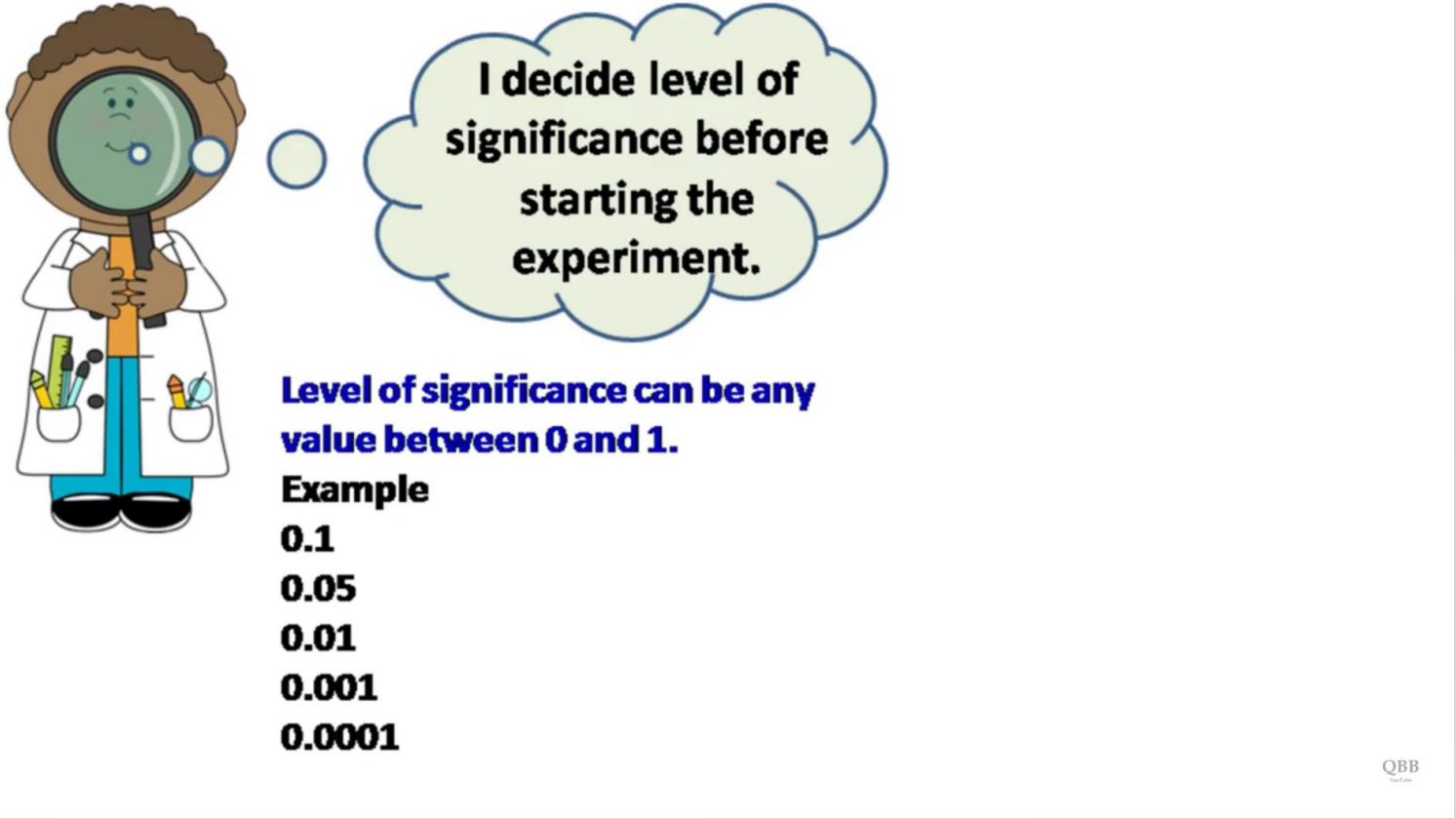


LEVEL OF SIGNIFICANCE

- The probability with which we reject a null hypothesis when its true is the level of significance α
- The probability with which we accept the null hypothesis when its true is called confidence interval $1-\alpha$



THE RESEARCHER FIXES THE LEVEL OF SIGNIFICANCE BEFORE STARTING THE EXPERIMENT

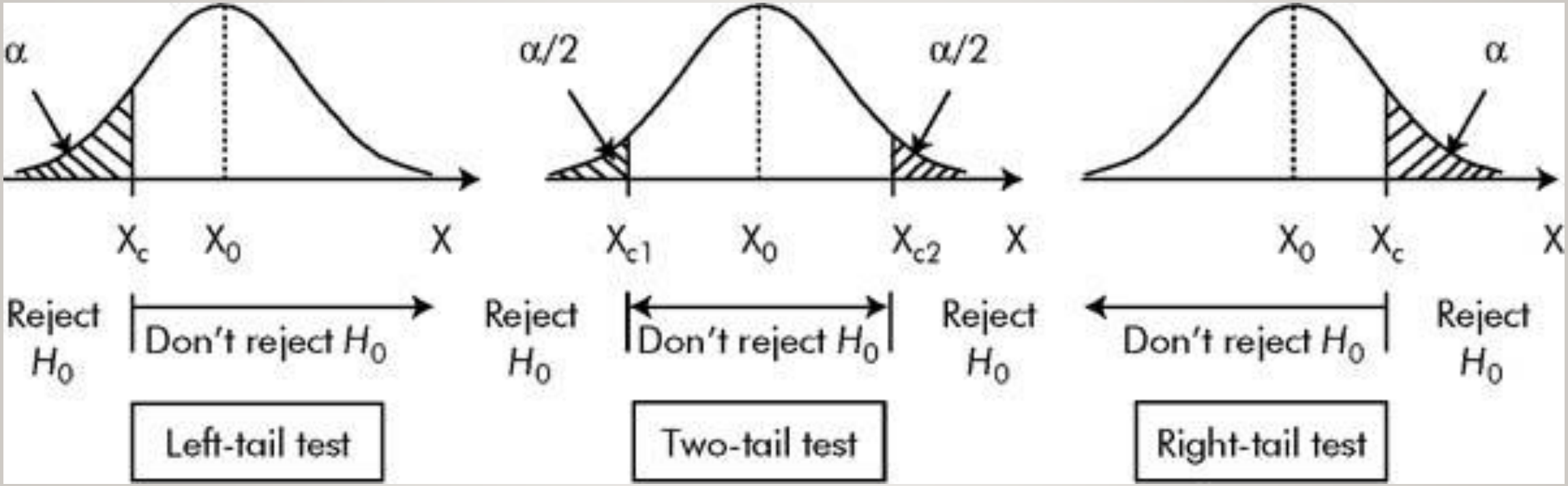


HYPOTHESIS TESTING STEPS

1. State the Null Hypothesis (H0) and Alternate Hypothesis (H1)
2. Choose the Level of Significance
3. Find Critical Values
4. Find test Statistic
5. Draw your conclusion

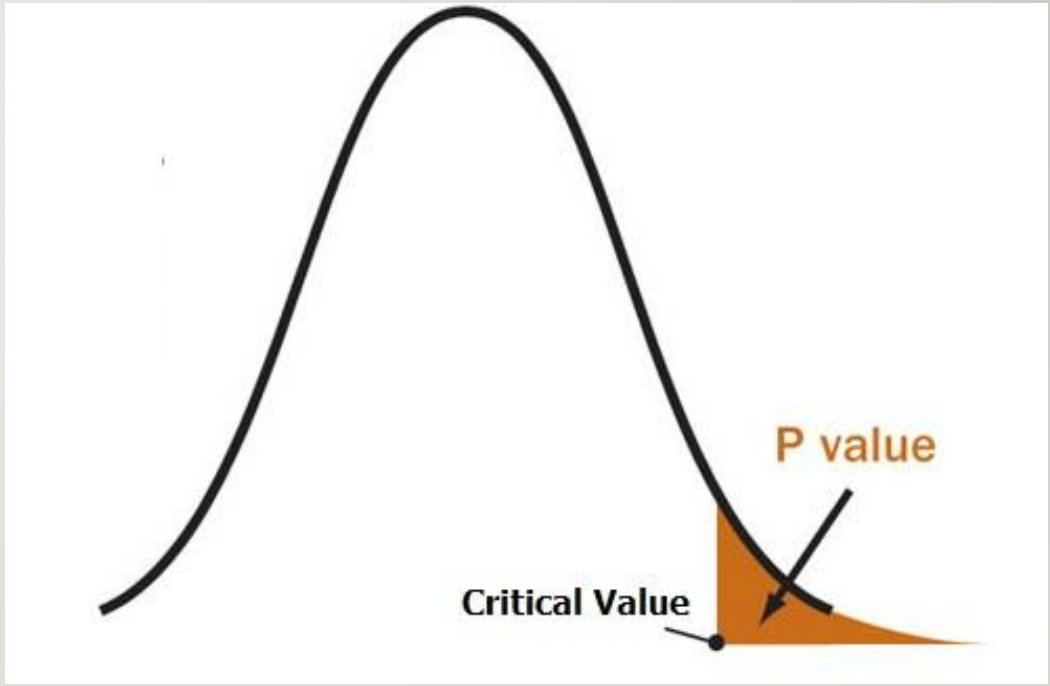
EXAMPLE HYPOTHESIS TESTING

- The average IQ for the adult population is 100 with a standard deviation of 15. A researcher believes this value has changed. The researcher decides to test the IQ of 75 random adults. The average IQ of the sample is 105. Is there enough evidence to suggest the average IQ has changed?

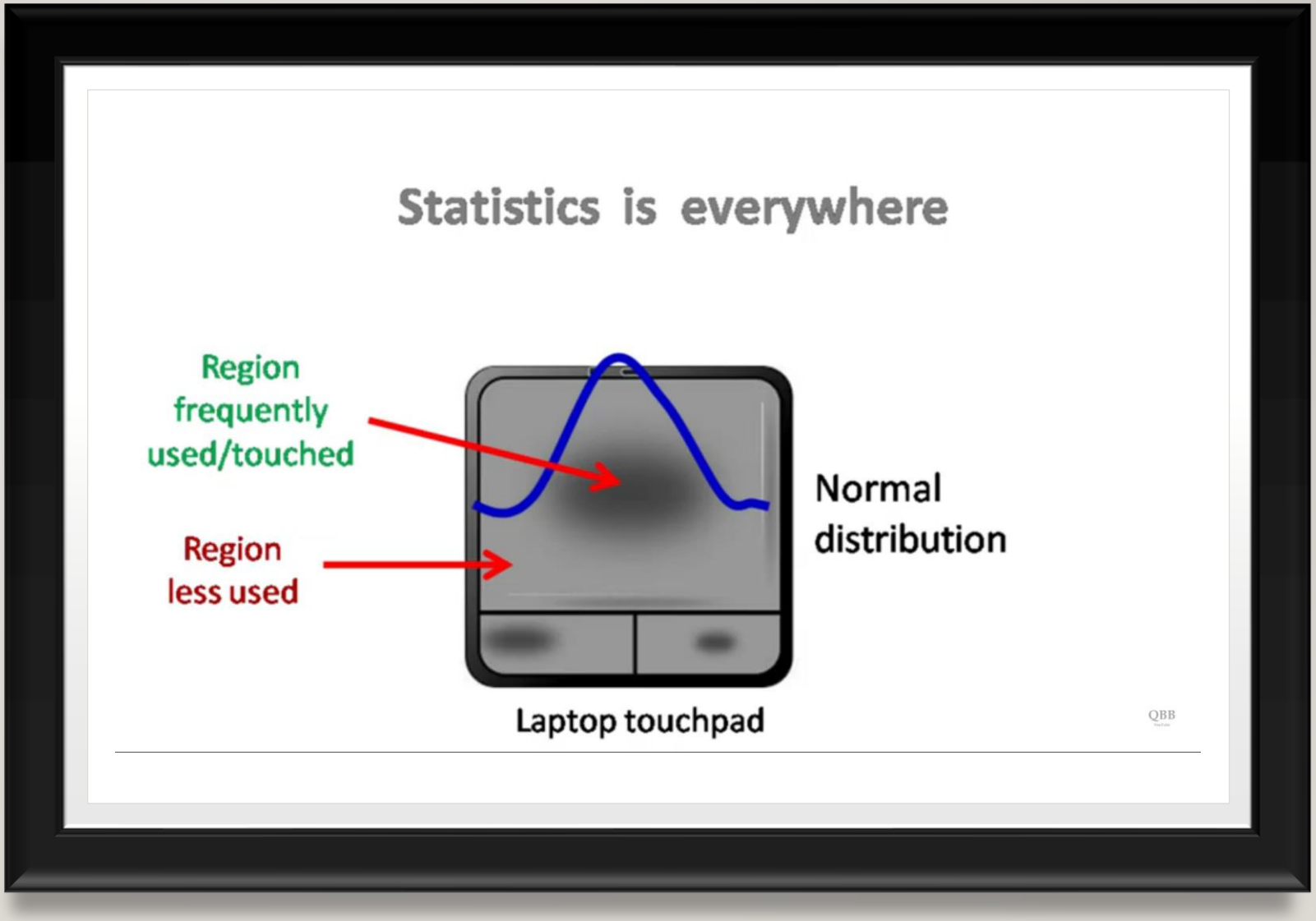

α
X_c
X_0
X
Reject H_0
Don't reject H_0
Left-tail test
α/2
α/2
X_{c1}
X_0
X_{c2}
X
Reject H_0
Don't reject H_0
Reject H_0
Two-tail test
α
X_0
X_c
X
Don't reject H_0
Reject H_0
Right-tail test

P-VALUE

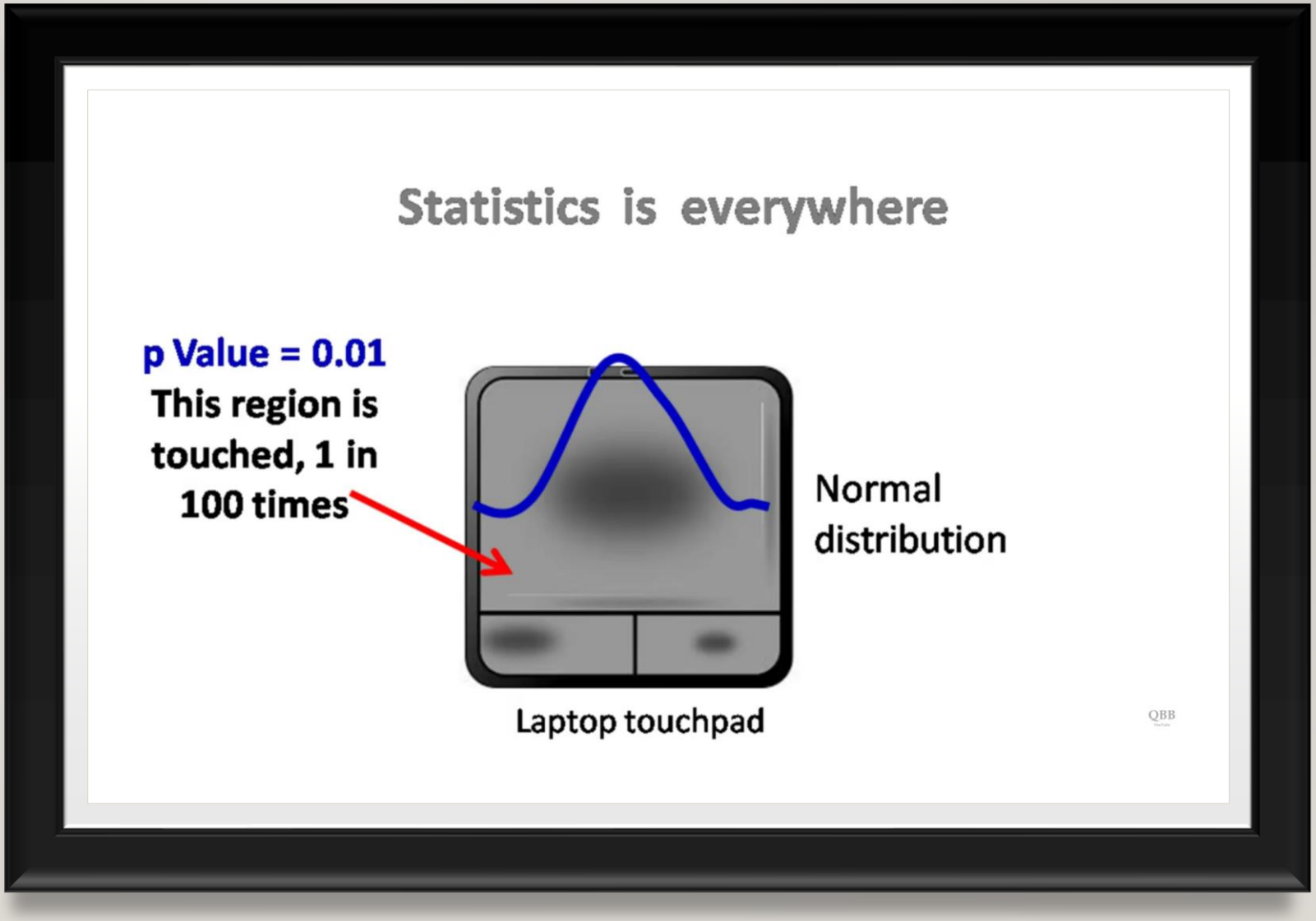
- A p-value helps you determine the significance of your results.
- The probability for the Null hypothesis to be true



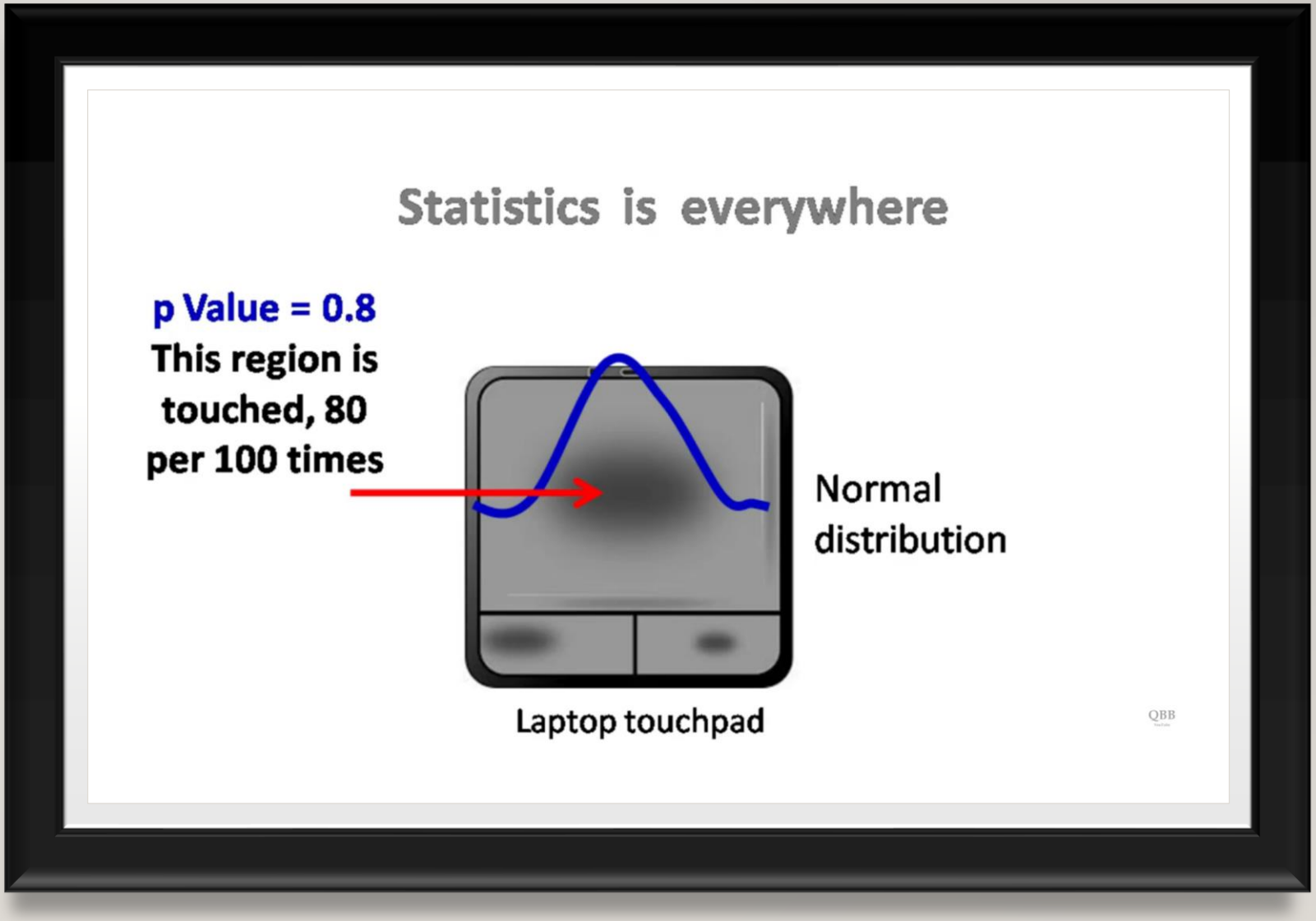
EXAMPLE FOR P - VALUE



EXAMPLE FOR P-VALUE



EXAMPLE FOR P-VALUE



P value (any value from 0 to 1)

P = 0.1

10 out of 100, null hypothesis will be true

P = 0.3

30 out of 100, null hypothesis will be true

P = 0.05

5 out of 100, null hypothesis will be true

P-VALUE

- Compare the p-value with α
  - If p-value $<$ α , reject H_0
  - If p-value $\geq$ α , do not reject H_0

EXAMPLE P-VALUE METHOD

- The average IQ for the adult population is 100 with a standard deviation of 15. A researcher believes this value has changed. The researcher decides to test the IQ of 75 random adults. The average IQ of the sample is 105. Is there enough evidence to suggest the average IQ has changed?

EXAMPLE FOR P- VALUE

- A company is looking to improve their website performance. Currently pages have a mean load time of 3.125 seconds, with a standard deviation of 0.700 seconds. They hire a consulting firm to improve load times. Management wants a 99% confidence level. A sample run of 40 of the new pages has a mean load time of 2.875 seconds. Are these results statistically faster than before?

EXAMPLE FOR P- VALUE

- A company is looking to improve their website performance. Currently pages have a mean load time of 3.125 seconds, with a standard deviation of 0.700 seconds. They hire a consulting firm to improve load times. Management wants a 99% confidence level. A sample run of 40 of the new pages has a mean load time of 2.875 seconds. Are these results statistically faster than before?

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

$$Z = -2.259$$
$$P = 0.0119$$

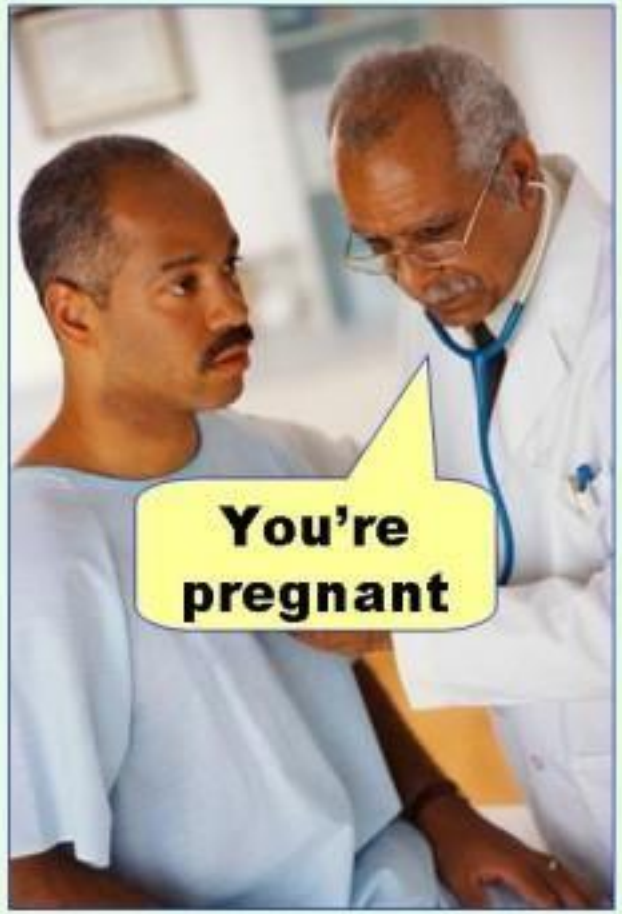
TYPE I AND TYPE II ERROR

- A **type I error** is the incorrect rejection of a true null hypothesis.
- A **type II error** is the failure to reject a false null hypothesis.

H_0: *Not pregnant*

H_1: *Are pregnant*

Type I error
(false positive)



Type II error
(false negative)