P Value

The Hypothesis Testing Modern Method

Our entire analysis can be boiled down to one number called the p-value. And, if that value (p-value) is less than the level of significance α, then we can say we have found statistical evidence in favor of the Alternative Hypothesis H_1. **Automatically accept the Alternate Hypothesis $\boldsymbol{H_1}$.**

$$\boldsymbol{p < \alpha}$$

Technical Note

The alternative hypothesis H_1 may, or may not be, the claim. So that means our conclusions are still either one of the following statements.

- The sample supports the claim.
- The sample does not support the claim.

It all depends on the p-value, the level of significance, and where your claim happens to be a Null Hypothesis or an Alternate Hypothesis.

The **p-value** tells us the probability of obtaining the test statistic we observed (based on our sample) assuming the null hypothesis H_0 is correct. The level of significance α is a borderline chance (probability) used for making a statistical decision in hypothesis testing.

$$\boldsymbol{\alpha = 1\%\ ,\ \alpha = 5\%,\ \text{or}\ \alpha = 10\%}$$

In summary, we are willing to take an α percent chance of rejecting our null hypothesis H_0 when it is actually true.

Type I and Type II Errors

When testing a hypothesis (claim) we arrive at a conclusion of either rejecting it or failing to reject it (accept it). Our conclusions are sometimes correct or sometimes wrong, even when we follow our procedure correctly. In essence, our hypothesis testing is based on likelihood and is not certain.

Type I Error- The mistake of rejecting the null hypothesis H_0 when it is true.

$$\alpha = p(\text{rejecting } H_0 \text{ when it is actually true})$$

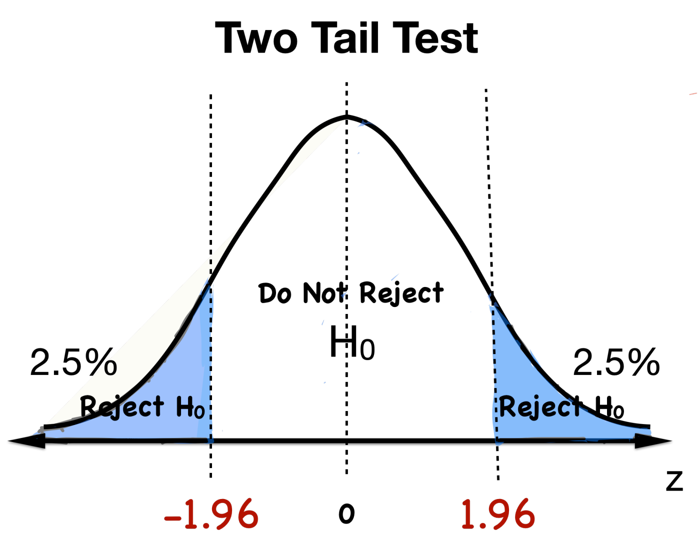
Type II Error - The mistake of failing to reject (accepting) the null hypothesis H_0 when it's false.

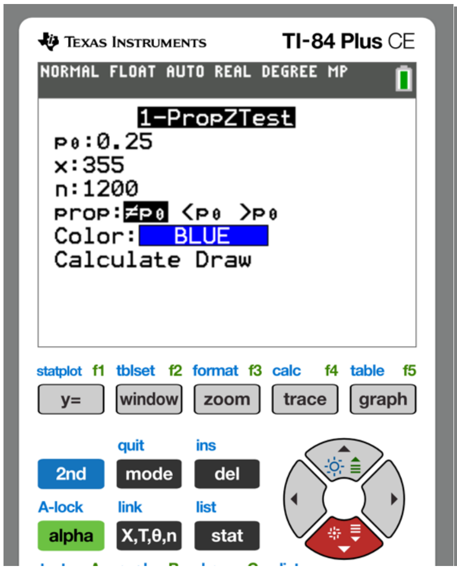
$$B = p(\text{accepting } H_0 \text{ when it is actually false})$$

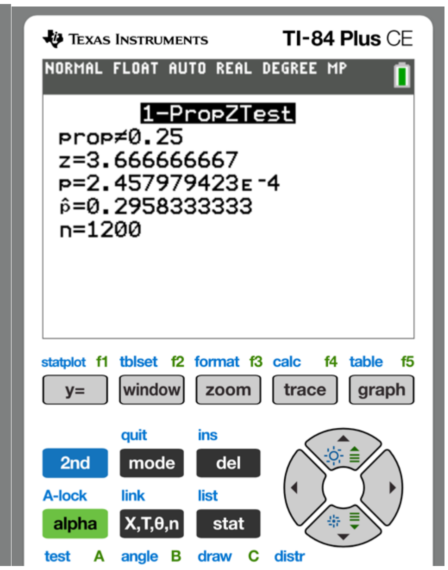
Example

The proportion of college students who were infected with covid 19 during the pandemic is 25% as claimed by Professor Snodgrass. A sample of 1200 college students reveal that 355 were infected with covid 19. Use the **5% level of significance** to test this claim using the **p-value method.**

$$H_0: p = 25\% \text{ claim}$$
$$H_1: p \neq 25\%$$







p-value and α

$$p \approx 2.45 \times 10^{-4} \approx \mathbf{0.000245} < \mathbf{5}\% = 0.05$$

We automatically accept the Alternate Hypothesis H_1

$H_0: p = 25\%$ **claim**
$\boldsymbol{H_1: p \neq 25\%}$

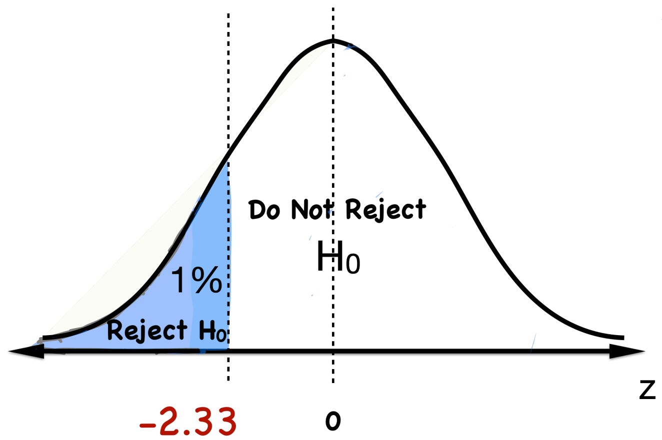
Conclusion: The sample does not support the claim.

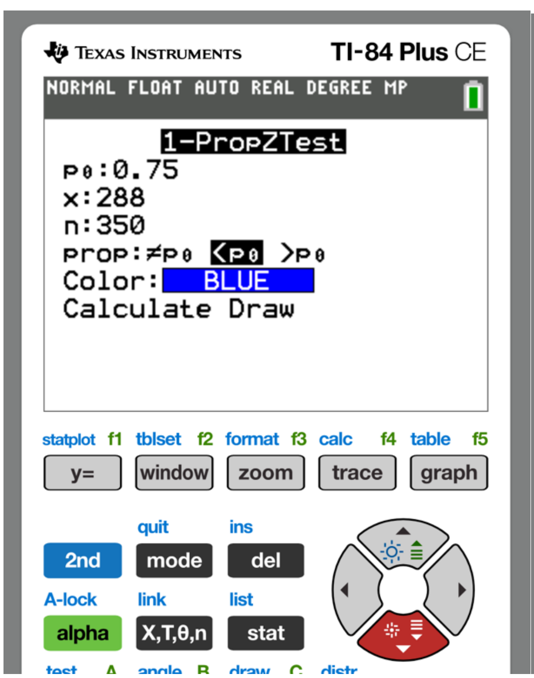
Example

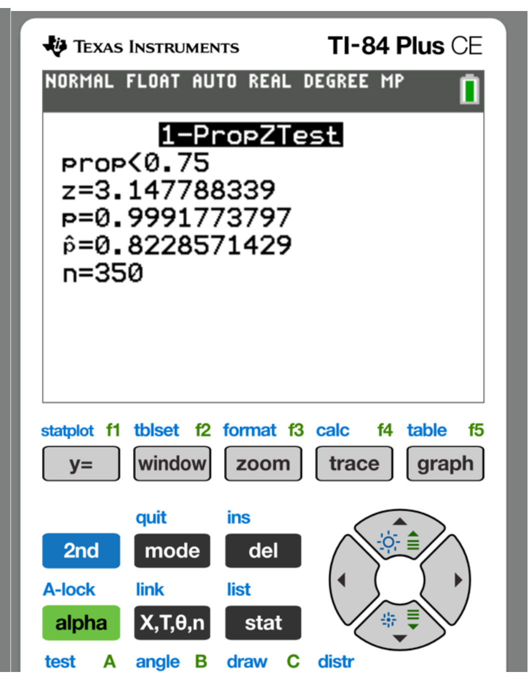
The proportion of college students who have a Netflix account is less than 75% as claimed by Professor Snodgrass. A sample of 350 college students reveal that 288 have a Netflix account. Use the **1% level of significance** to test this claim by **the p-value method**.

$$H_0\colon p \geq 75\%$$
$$H_1\colon p < 75\% \quad \textbf{claim}$$

Left Tail Test







p-value and α

$$p \approx 0.999 \nless \mathbf{1}\% = 0.01$$

We do not automatically accept the Alternate Hypothesis $\boldsymbol{H_1}$
We accept the Null Hypothesis $\boldsymbol{H_0}$

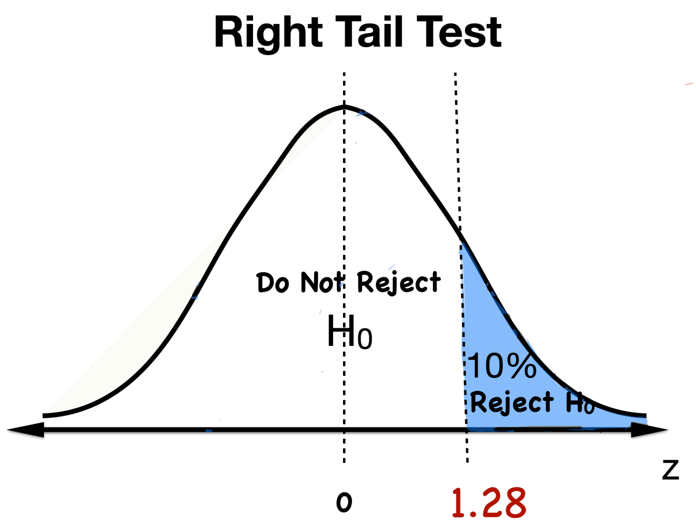
$$H_0\colon p \geq 75\%$$
$$H_1\colon p < 75\%$$ **claim**

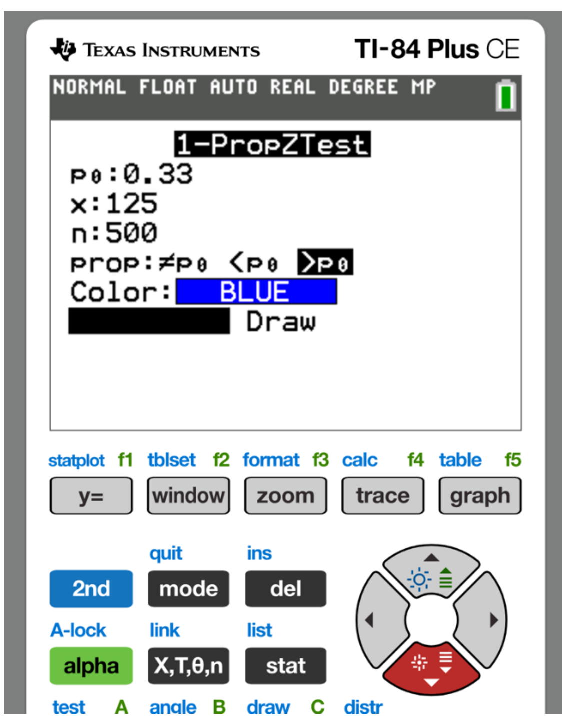
Conclusion: The sample does not support the claim.

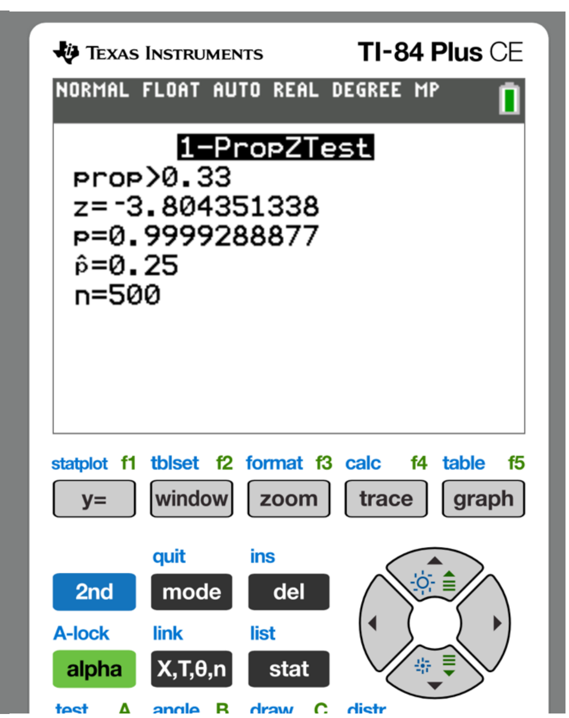
Example

The proportion of college students who get a good night's sleep (at least 8 hours) the night before a final exam is no more than 33% as claimed by Professor Snodgrass. A sample of 500 college students reveal that 125 got a good night's sleep (at least 8 hours). Use the **10% level of significance** to test the hypothesis by the **p-value method**.

$$H_0: p \leq 33\%$$ **claim**
$$H_1: p > 33\%$$







p-value and α

$$p \approx 0.999 \nless \mathbf{10}\% = 0.10$$

We do not automatically accept the Alternate Hypothesis $\boldsymbol{H_1}$
We accept the Null Hypothesis $\boldsymbol{H_0}$

$H_0: p \leq 33\%$ **claim**
$H_1: p > 33\%$

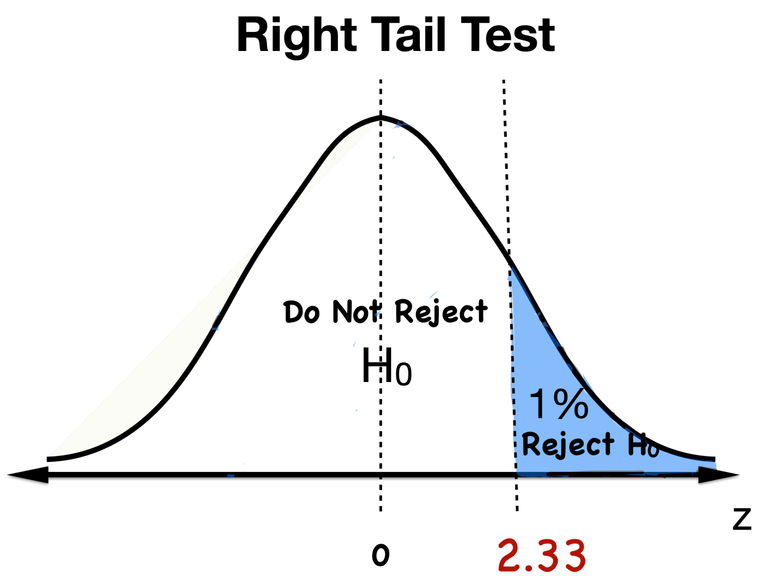
Conclusion: The sample supports the claim.

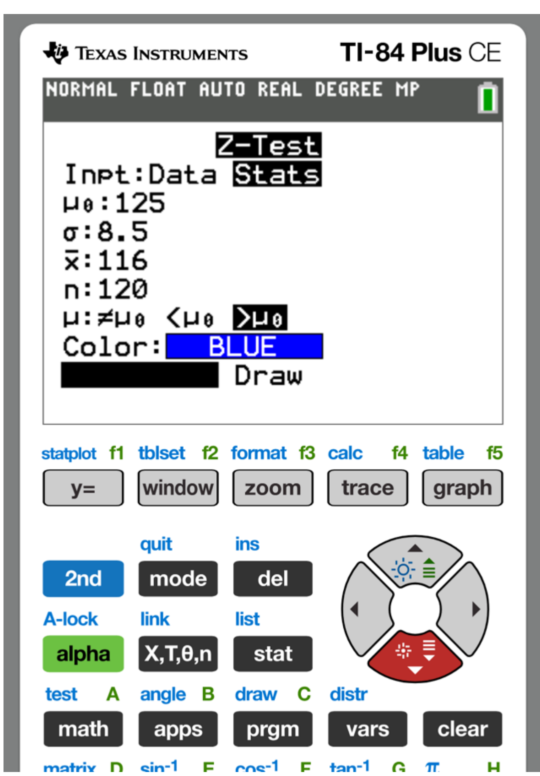
Example

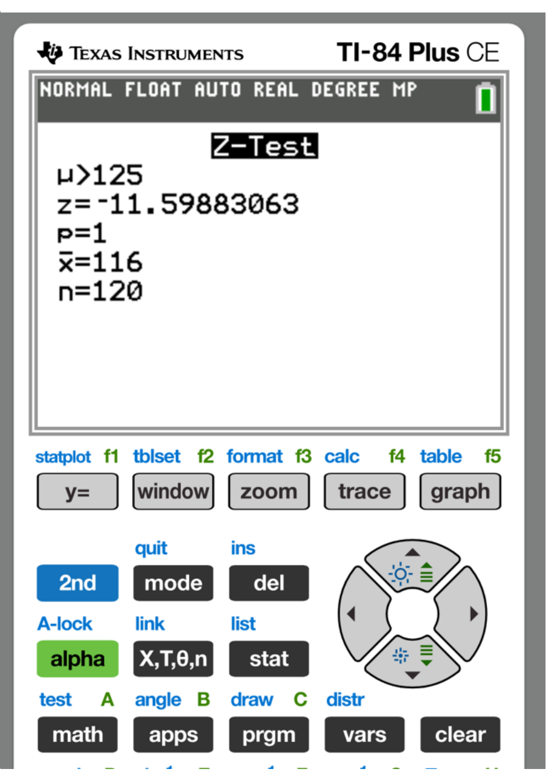
The mean IQ scores for College Professors is no more than 125 as claimed by the Faculty Association. A sample of 120 College Professors reveal a mean IQ score of 116 with a standard deviation of 8.5. Use the 1% level of significance to test this claim.

$H_0: \mu \leq 125$ claim

$H_1: \mu > 125$







p-value and α

$$p = 1 \nless \mathbf{1}\% = 0.01$$

We do not automatically accept the Alternate Hypothesis $\boldsymbol{H_1}$
We accept the Null Hypothesis $\boldsymbol{H_0}$

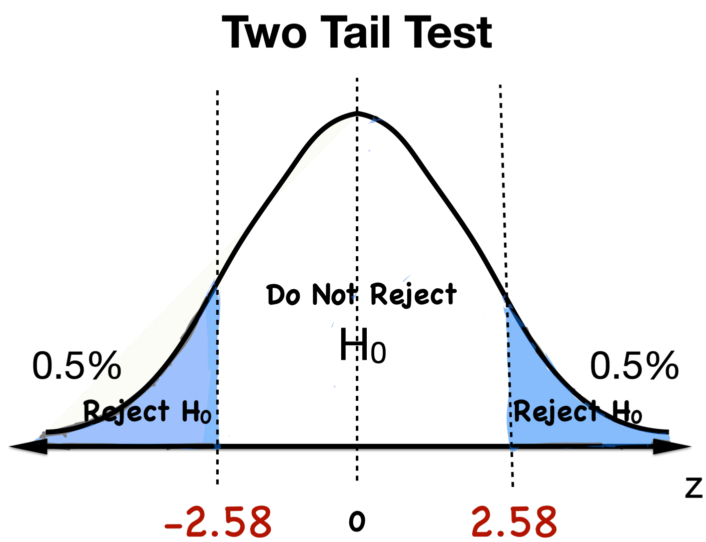
$H_0: p \leq 33\%$ **claim**
$H_1: p > 33\%$

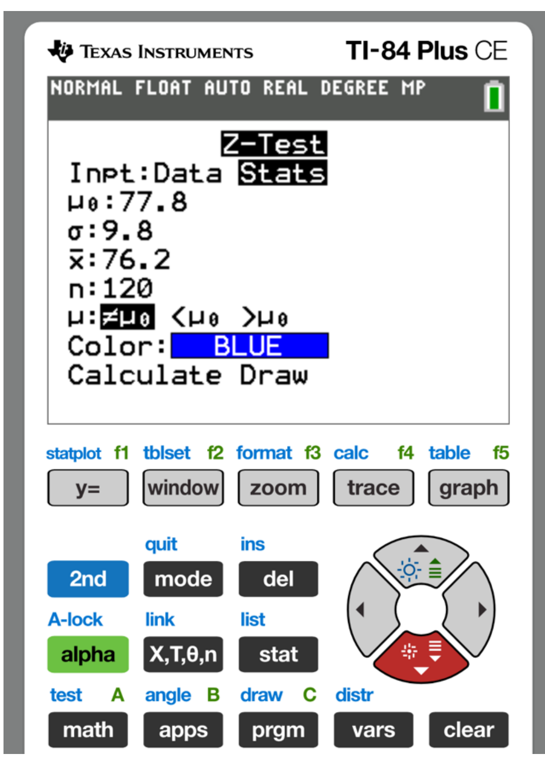
Conclusion: The sample supports the claim.

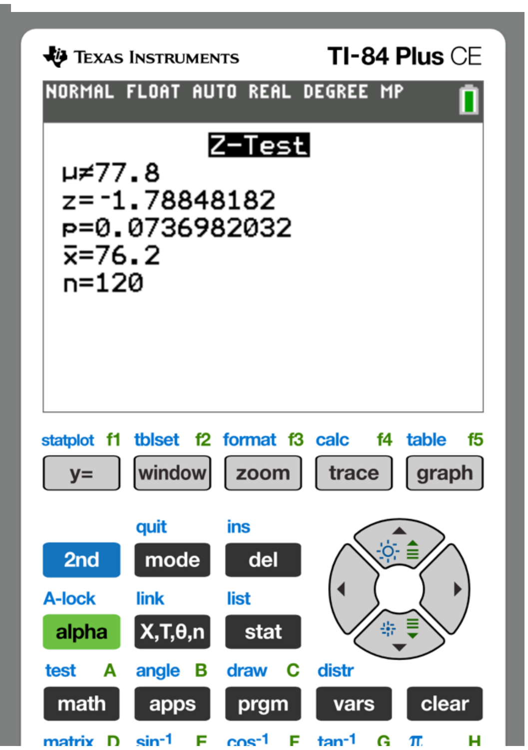
Example

The mean lifespan of California resident's is 77.8 years as claimed by the State of California. A sample of 50 California residents reveal a mean lifespan of 76.2 years with a standard deviation of 9.8 years. Use the 1% level of significance to test this claim.

$$H_0: \mu = 77.8 \text{ claim}$$
$$H_1: \mu \neq 77.8$$







p-value and α

$$p \approx 0.074 \nless \mathbf{1}\% = 0.01$$

We do not automatically accept the Alternate Hypothesis $\boldsymbol{H_1}$
We accept the Null Hypothesis $\boldsymbol{H_0}$

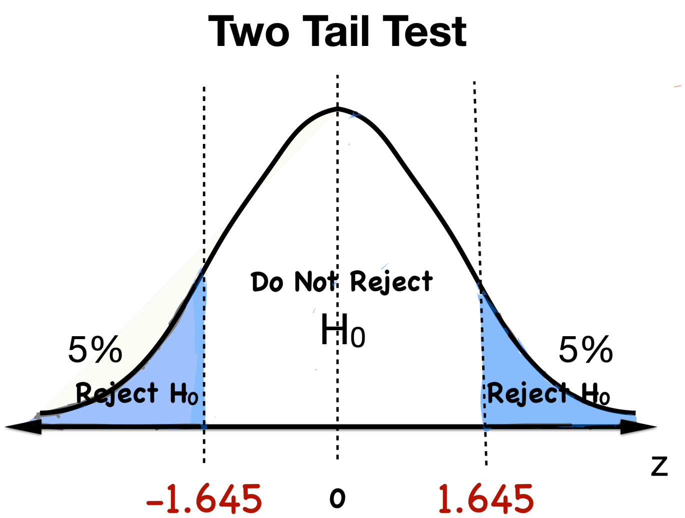
$H_0: \mu = 77.8$ **claim**
$H_1: \mu \neq 77.8$

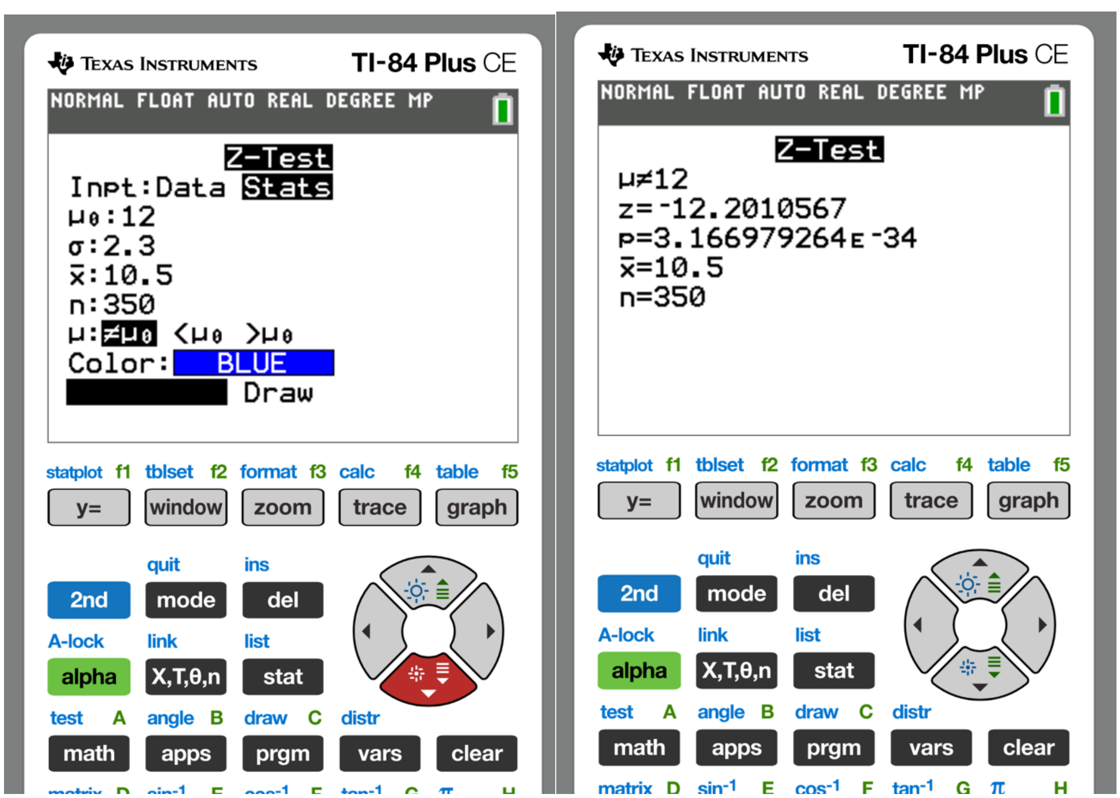
Conclusion: The sample supports the claim.

Example

The mean number of units college students take in a remote (online) learning environment is not 12 units as claimed by campus researchers. A sample of 350 college students reveal a mean of 10.5 units with a standard deviation of 2.3 units. Use the 10% level of significance to test this claim.

$$H_0: \mu = 12$$
$$H_1: \mu \neq 12 \text{ claim}$$





p-value and α

$$p \approx 0.000 < \mathbf{10}\% = 0.10$$

We automatically accept the Alternate Hypothesis $\boldsymbol{H_1}$

$H_0: \mu = 12$
$H_1: \mu \neq 12$ **claim**

Conclusion: The sample supports the claim.