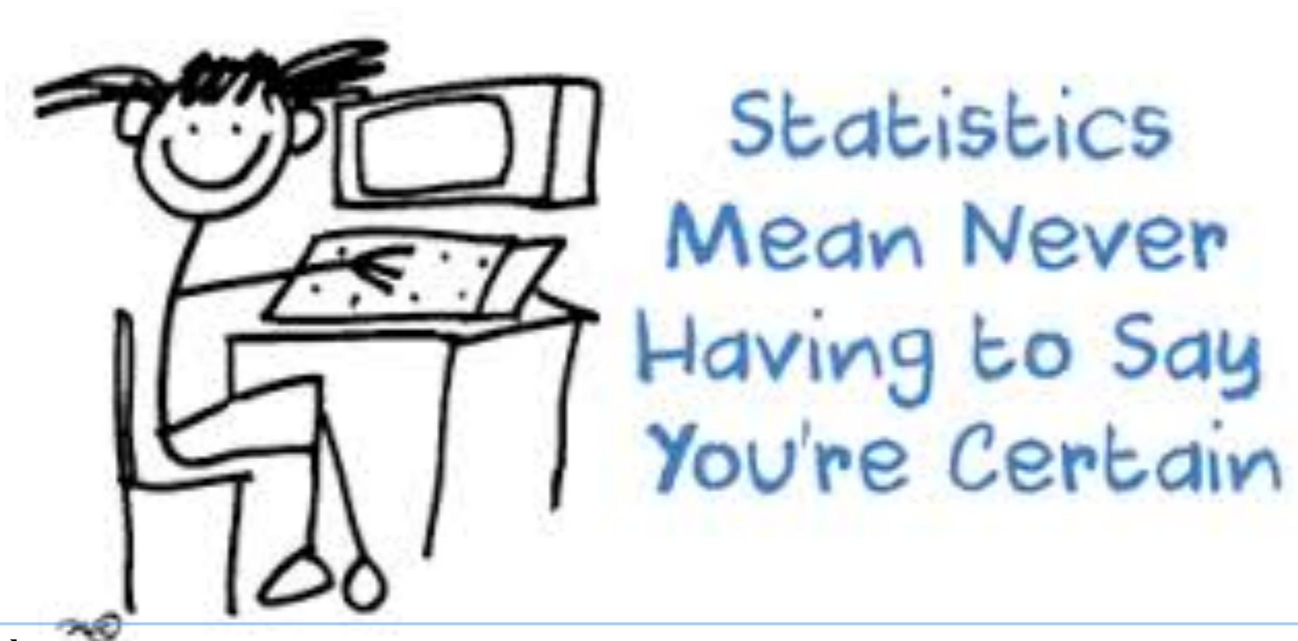
Hypothesis Testing



Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ?

mth 150

Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ? **or** Can you be 95% confident that the new therapy does cahge the mean survival time?

$n = 100$ Confidence = 0.95
$x_o = 37$
sigma = 6.5
Margin of error, E = 1.273

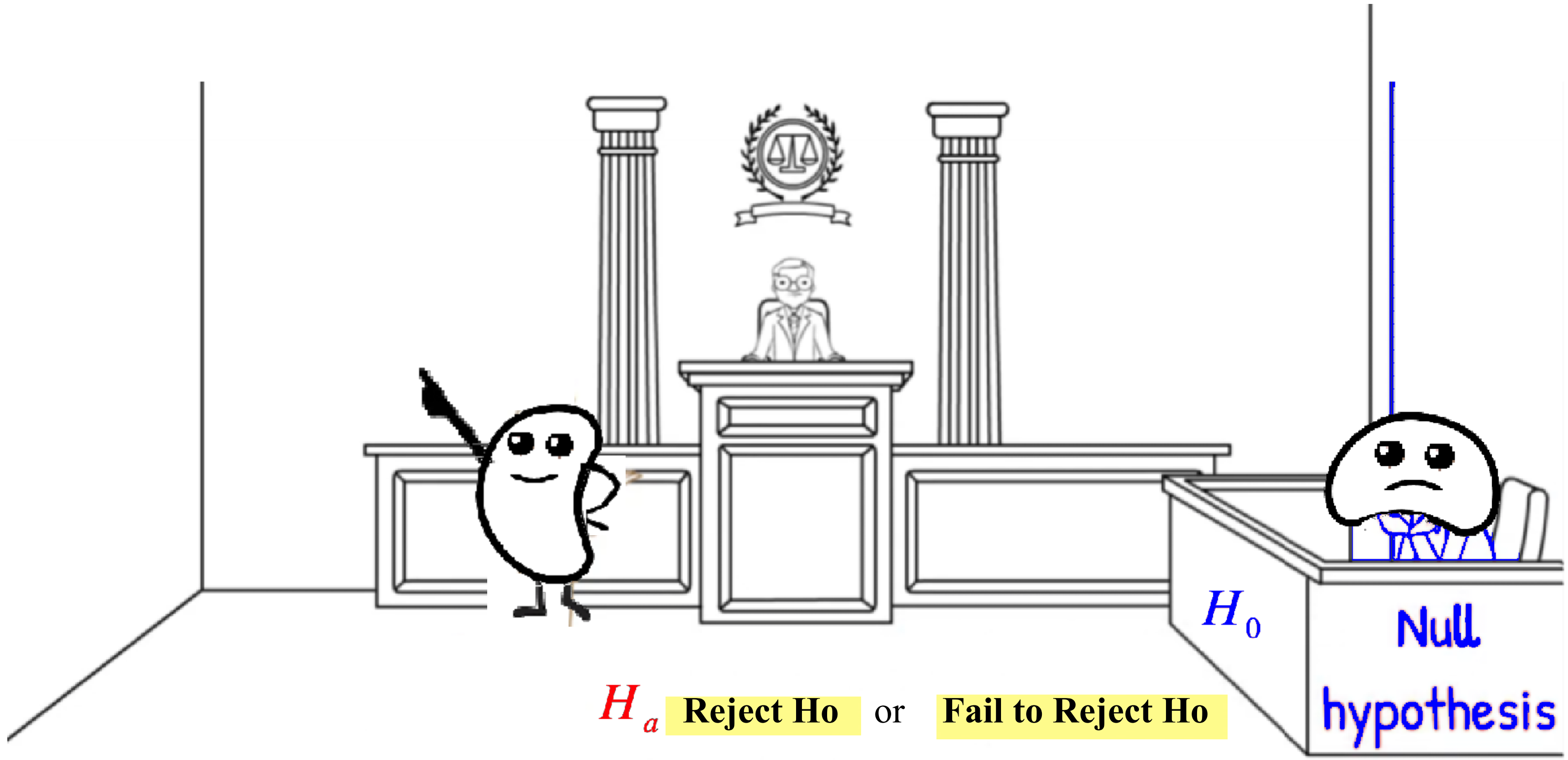
95% Confident the population mean is within the range:

$35.726 < \text{mean} < 38.273$

If our confidence interval contains the value claimed by the null hypothesis, then our sample result is close enough to the claimed value, and we therefore do not reject H_0.

If our confidence interval does not contain the value claimed by the null hypothesis, then our sample result is different enough from the claimed value, and we therefore reject H_0.

The Null Hypothesis



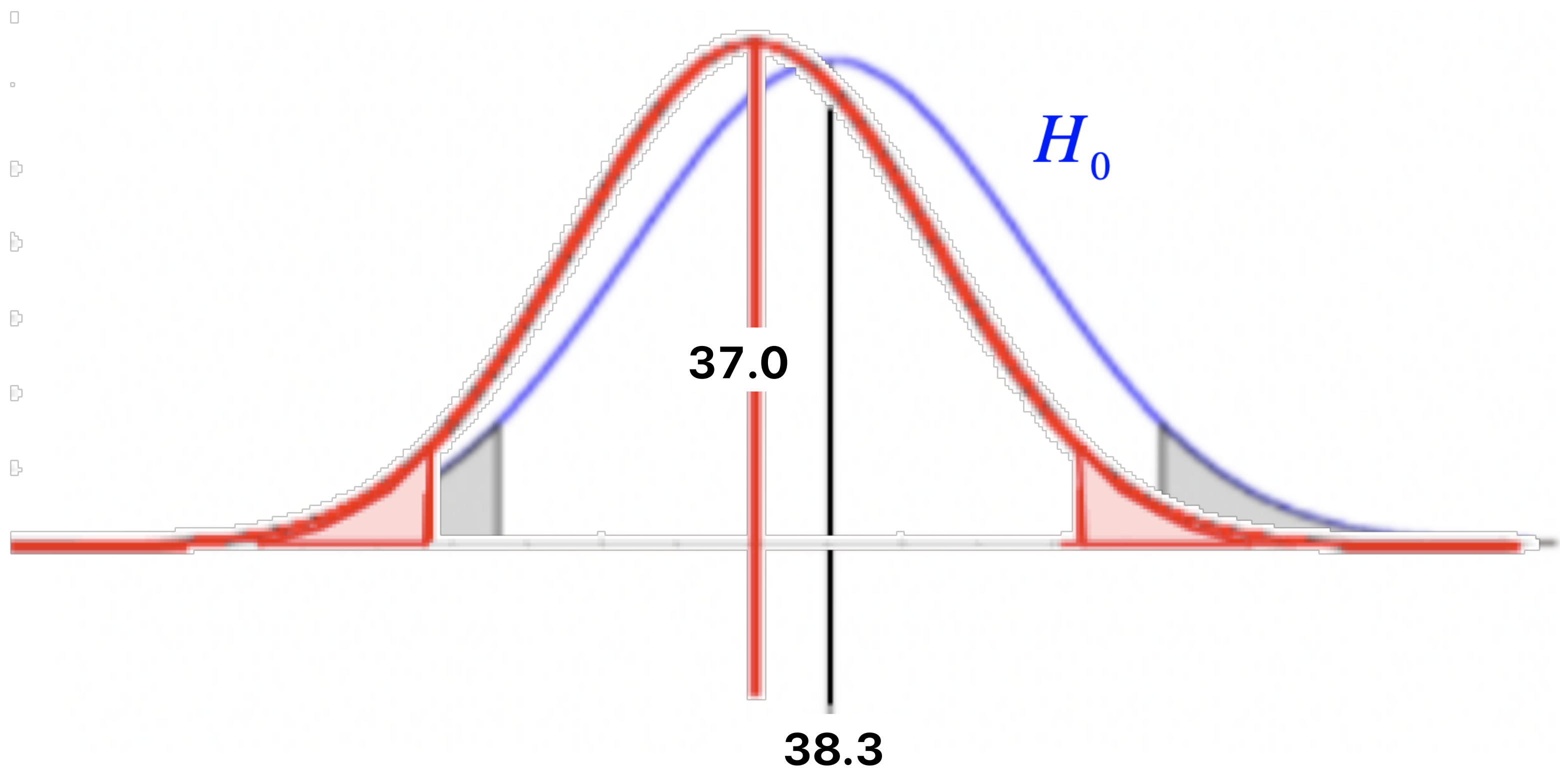
Why Dont Statisticians ACCEPT the Null Hypothesis?

To understand why we dont accept the null, consider the fact that you cant prove a negative. A lack of evidence only means that you havent proven that something exists. It does not prove that something doesnt exist. It might exist, but your study missed it.

Example:

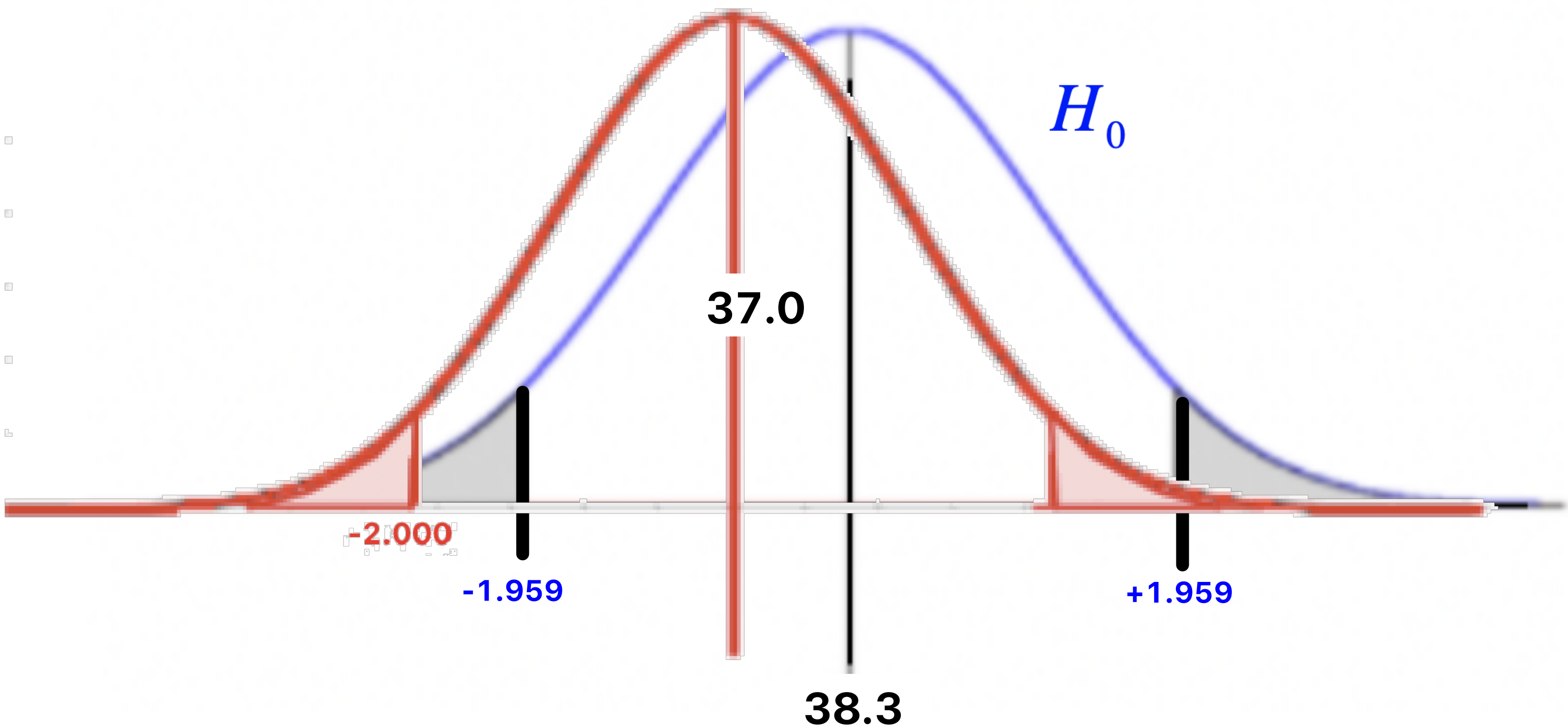
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ?

Two Tailed Test



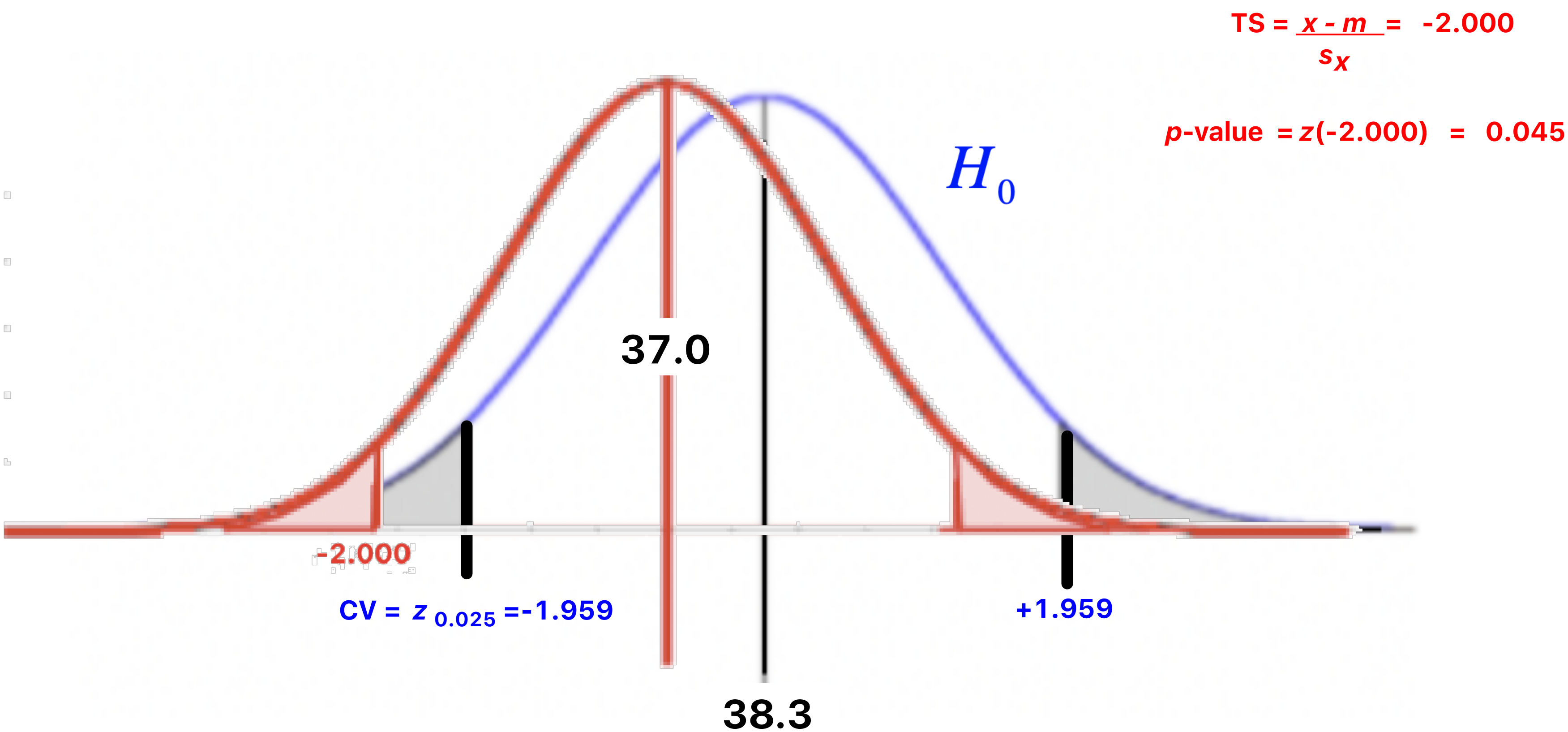
Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ?

Two Tailed Test

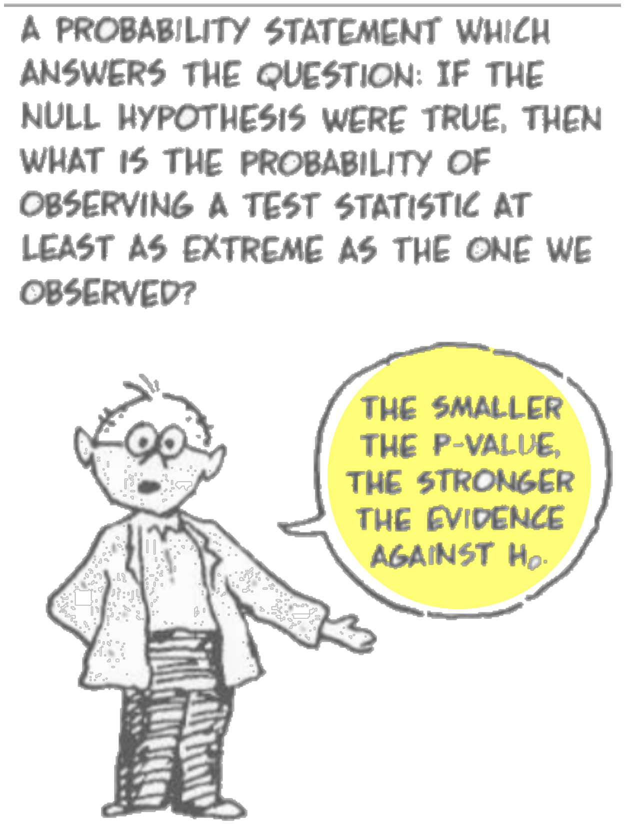


Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ?

Two Tailed Test



Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ?



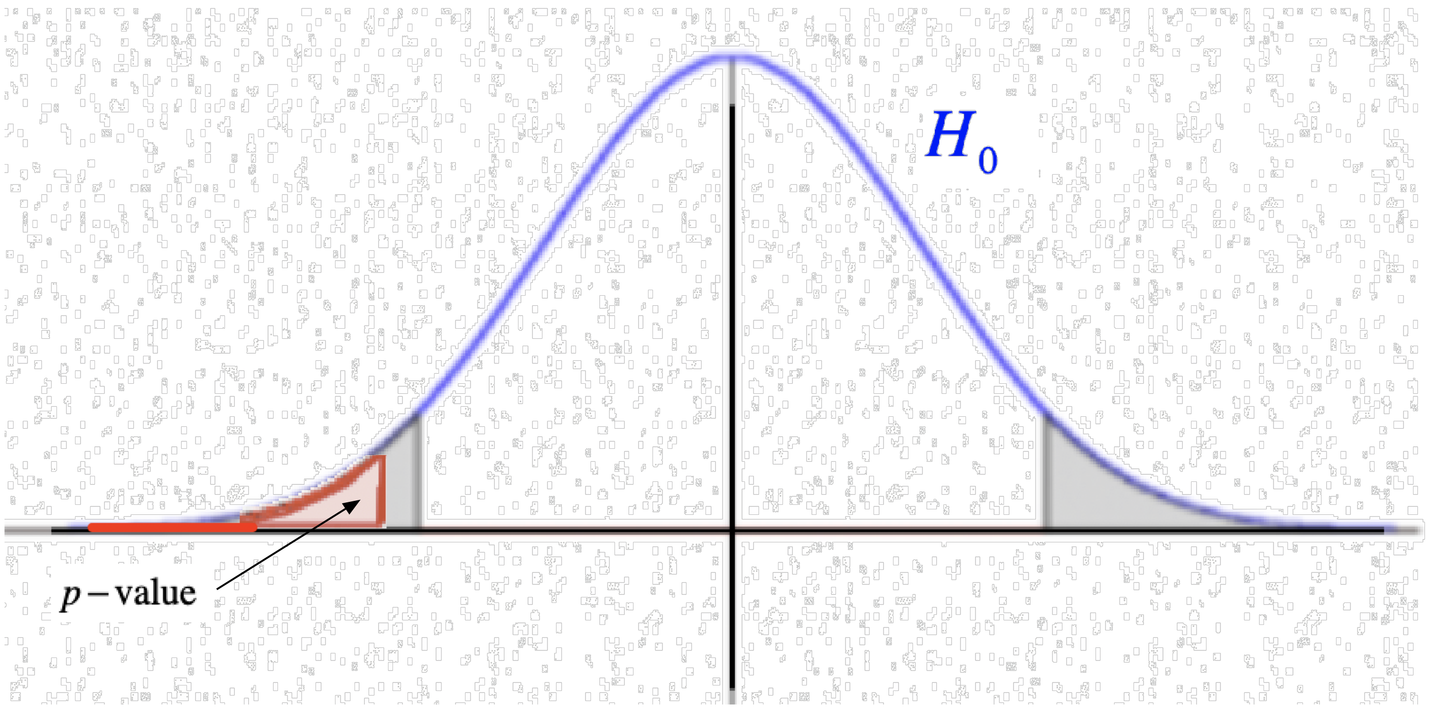
p-value $\leq \alpha$ -----> REJECT H_0

The ***p*-value** is the area in the appropriate tail(s) of the distribution of the test statistic (TS) when H_0 is true. That is the p-value will be

$$z_{p\text{ - value}} = 2.10 \Rightarrow p\text{-value} = 0.0179$$

$$p\left(\text{observing a test statistic as extreme or more extreme than we have} \mid H_0 \text{ is true}\right) = p\text{-value}$$

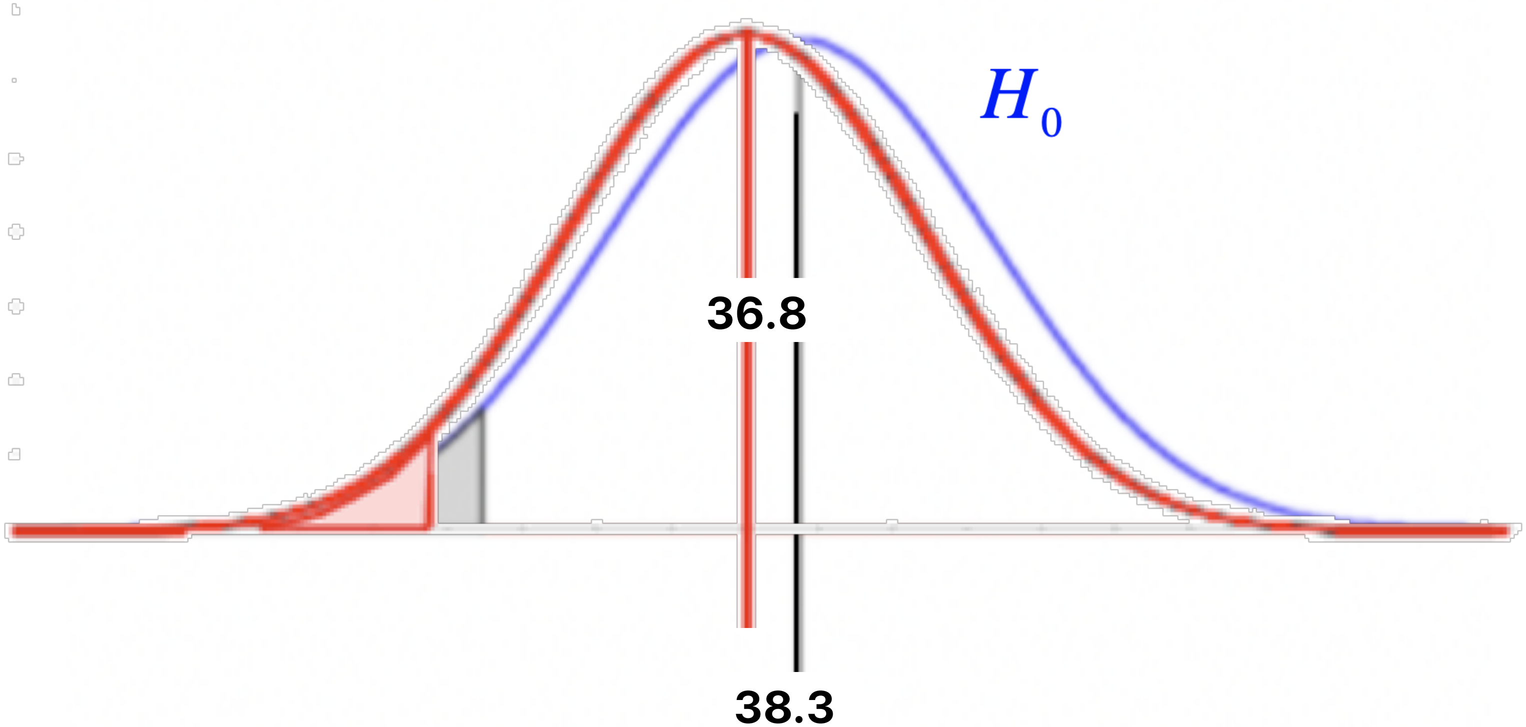
Two Tailed Test



38.3

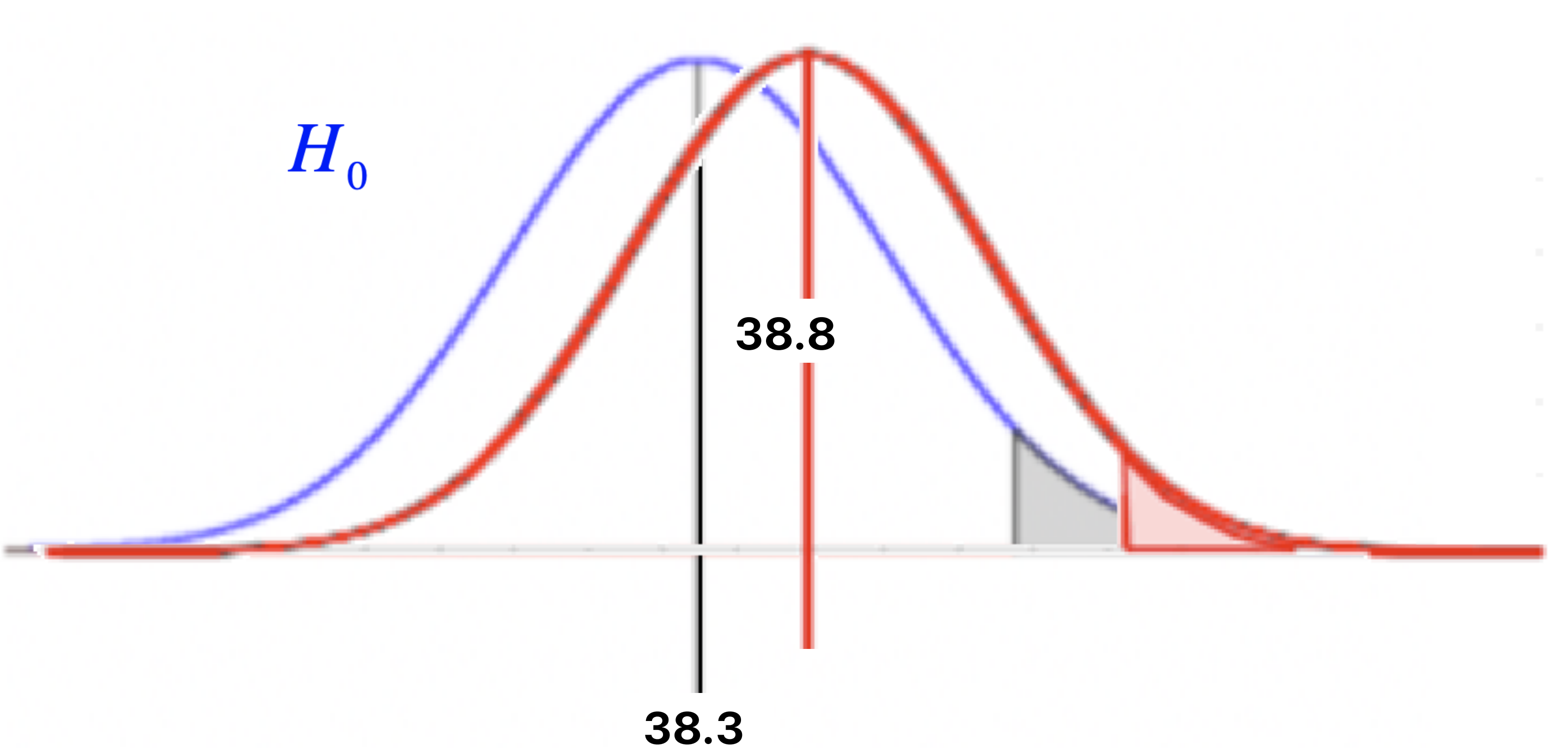
Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 37 months. Is survival statistically significantly changed (5%) with receipt of the new treatment ?

Left Tailed Test



Example:
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 36.8 months. Is survival statistically significantly lowered (5%) with receipt of the new treatment ?

Right Tailed Test



Example:

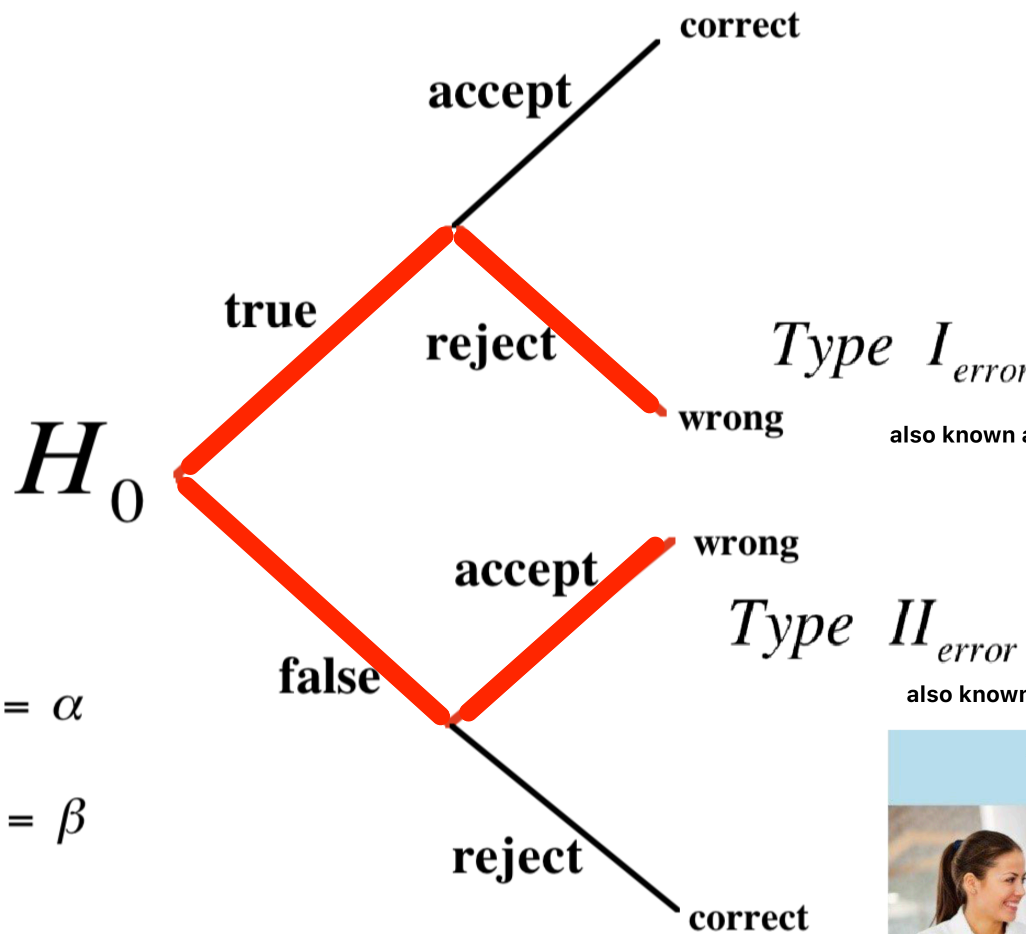
A virus has attacked the bean population and is expected to survive a mean duration of time equal to 38.3 months with a standard deviation of 6.5 months. Investigators are hopeful that a new therapy will affect survival. Suppose that the new treatment is administered to 100 cultures. It is observed that the average survival time is 38.8 months. Is survival statistically significantly increased (5%) with receipt of the new treatment ?

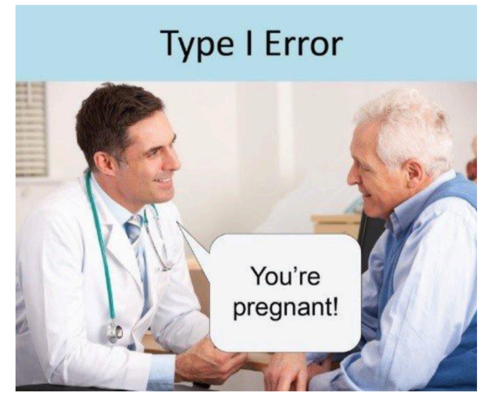
Is this increase statistically significant, or is it likely to be simply the result of chance variation?

Hypothesis Testing

- Technically, H_0 is constructed as the opposite of H_a .
- Hypothesis testing is just a probability calculation based on real world data.
- Hypothesis testing doesn't PROVE anything - it assesses, by counterexample, if it seems to be likely that H_0 is false.

Hypothesis Testing
(Type I and Type II errors)

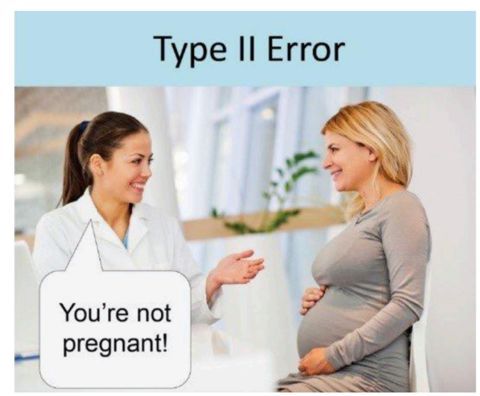




also known as a "false positive" finding or conclusion;

H_0 : You are NOT pregnant .

also known as a "false negative" finding or conclusion;

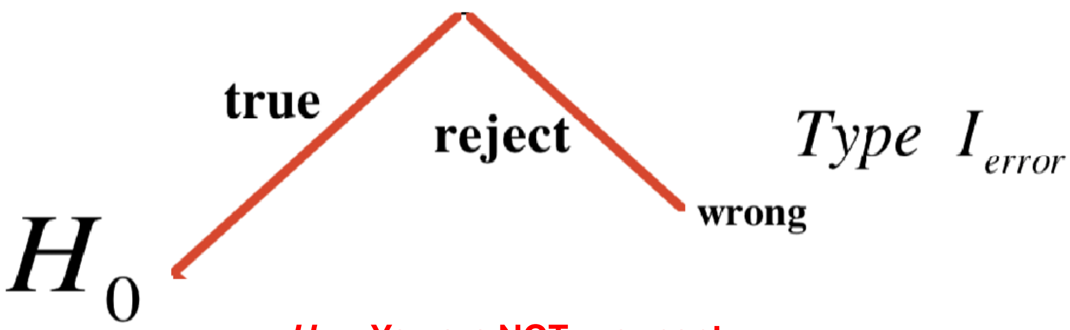


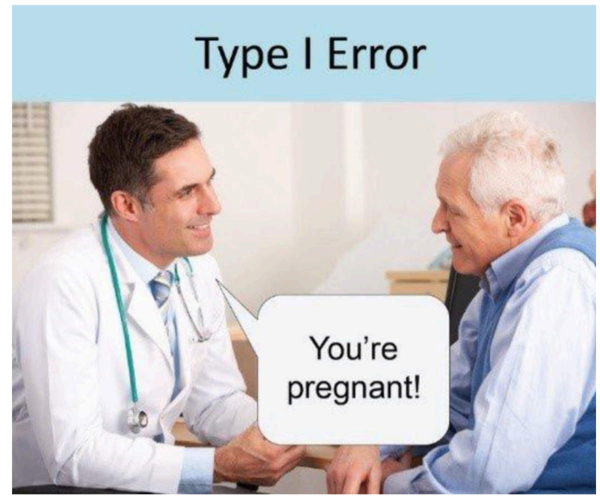
- $p(Type\ I_{error}) = \alpha$
- $p(Type\ II_{error}) = \beta$

Hypothesis Testing

- $p(Type\ I_{error}) = \alpha$

$$p(\text{incorrectly rejecting } H_0 \mid H_0 \text{ is true}) = \alpha$$





H_O : You are NOT pregnant .

"false positive"

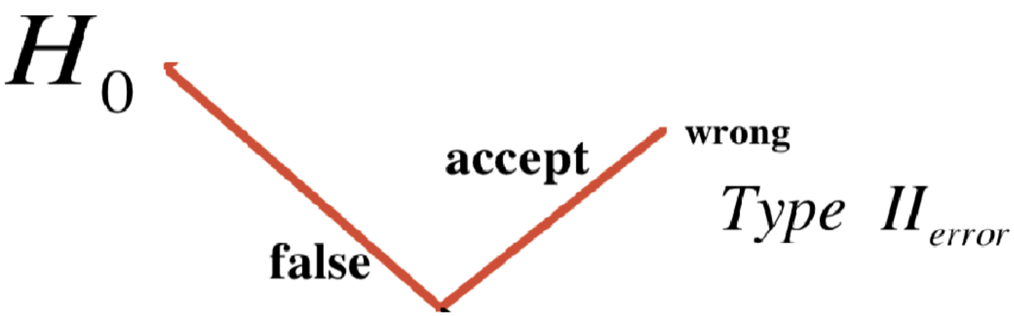
- **Set by the researcher and is based on how serious the consequences of the situation are.**

Hypothesis Testing

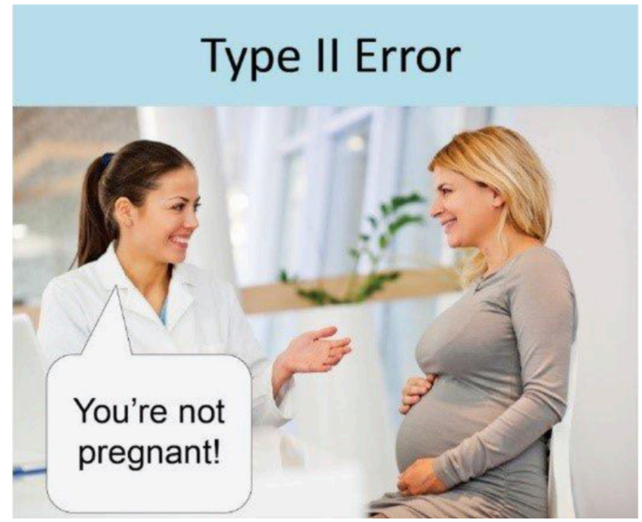
- $p(Type\ II_{error}) = \beta$

$$p(\text{incorrectly failing to reject } H_0 \mid H_0 \text{ is false}) = \beta$$

"false negative"

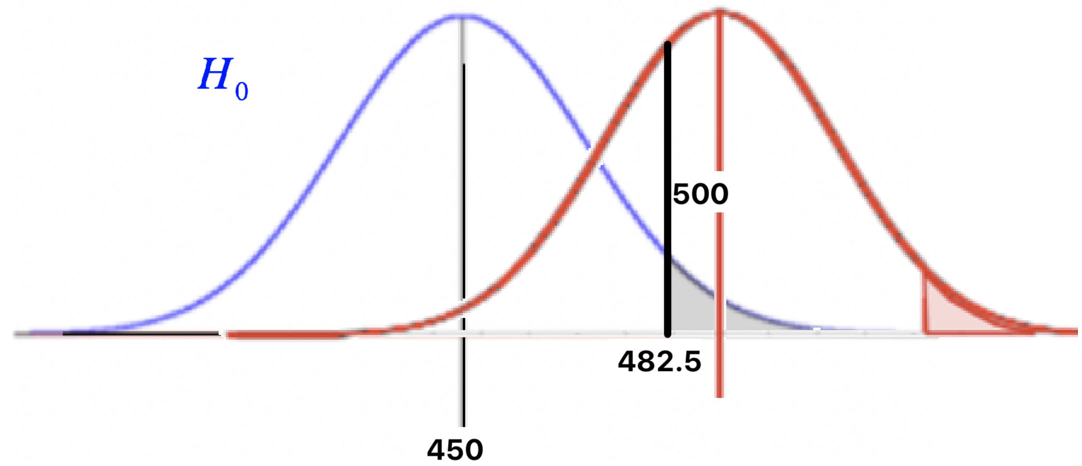


H_0 : You are NOT pregnant .



- $$p\left(\frac{\left[\bar{x} + t_{\nu,\,\alpha}\left(\frac{s}{\sqrt{n}} \right) \right] - \mu_0}{\frac{s}{\sqrt{n}}} \right)$$

Example:



For a mean of 450, standard deviation of 60, sample size of 9, $\bar{x} = 500$ and significance 0.05. The critical x value for a test statistic with $\alpha = 0.05$ would be

$$x = \bar{x} + z_{0.05}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$= 450 + 1.645\left(\frac{60}{\sqrt{9}}\right) = 482.9$$

β would be the probability associated with our TS, that is,

$$p(z \text{ is } TS)$$

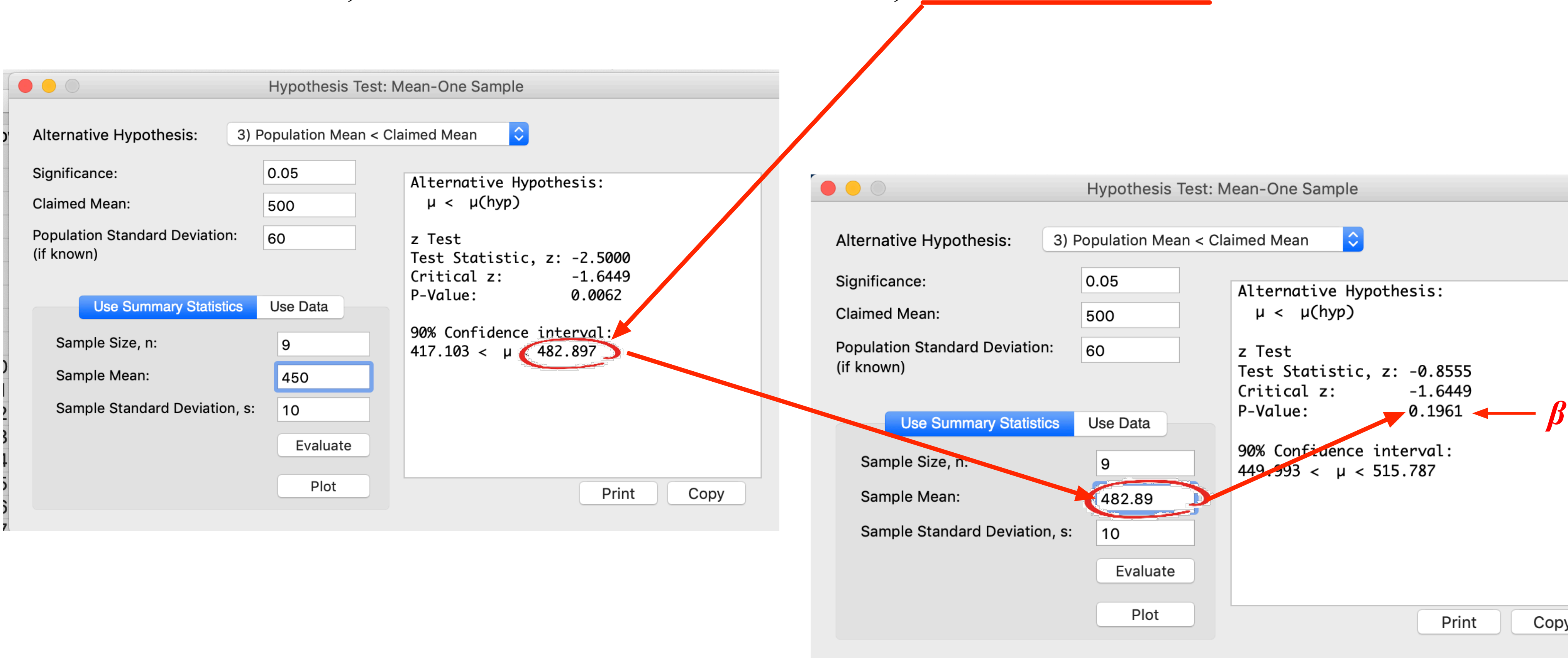
$$p\left(z \text{ is } \frac{482.9 - 500}{\frac{60}{\sqrt{9}}}\right) = p(z \text{ is } 0.8551) = 0.196$$

NOTE: to decrease β you would have to increase α. The only way to decrease both α and β would be to increase the sample size!

- See - Errors_and_Power.xls

Example: Calculating the probability of a Type II error.

- **For a mean of 450, standard devation of 60, sample size of 9, with claimed mean 500 and $\alpha = 0.05$, the critical x-value is**



NOTE: to decrease β you would have to increase α . The only way to decrease both α and β would be to increase the sample size!

- See - Errors_and_Power.xls

Statistical power or the power of a hypothesis test is the probability that the test correctly rejects the null hypothesis.

- Power is the probability of rejecting the null hypothesis when, in fact, it is false.
- Power is the probability of making a correct decision (to reject the null hypothesis) when the null hypothesis is false.
- Power is the probability that a test of significance will pick up on an effect that is present.
- Power is the probability that a test of significance will detect a deviation from the null hypothesis, should such a deviation exist.
- Power is the probability of avoiding a Type II error.
- Power = 1 - β

$$p(\text{correctly rejecting } H_0 \mid H_0 \text{ is false}) = 1 - \beta$$

from the previous example, the power of our test would be 1 - β or 1 - 0.280 = 0.729