Confidence Intervals and Hypothesis Testing

Module 4e

Type 1 & 2 Errors and P-values

Learning Objectives

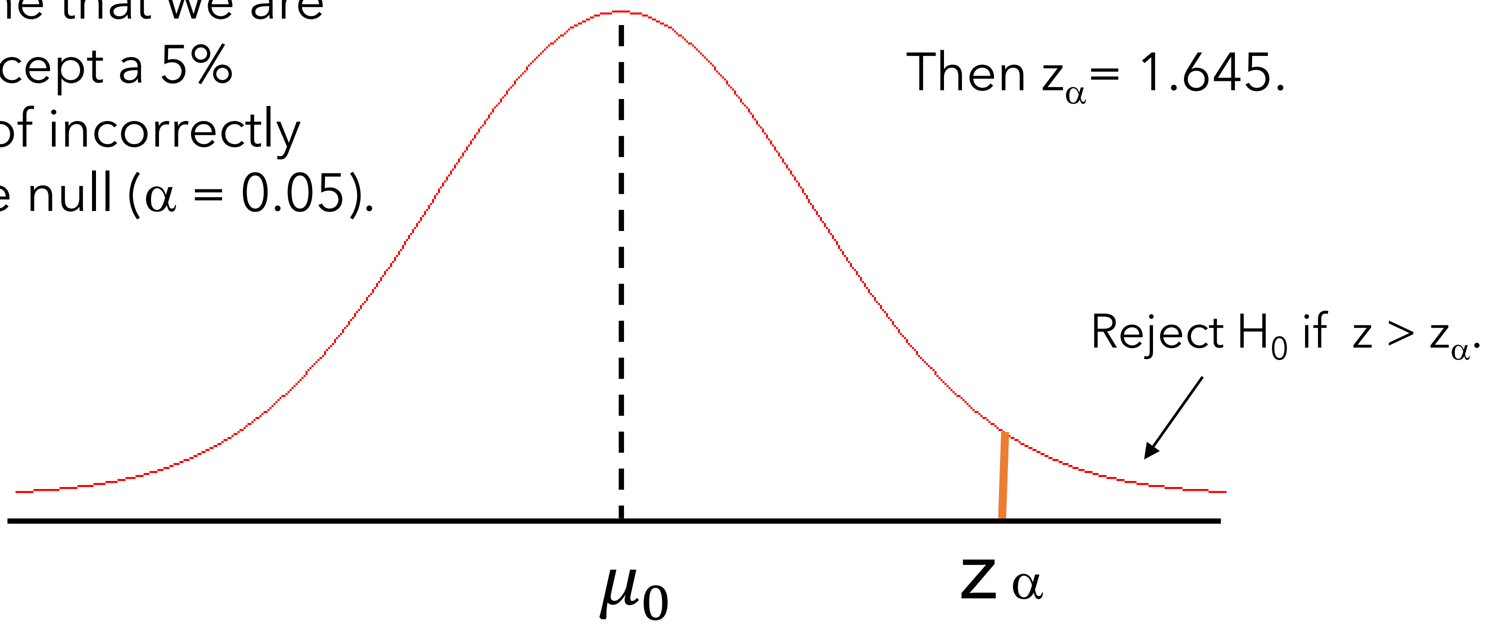
- Choose an appropriate significance level-- α
- Understand and balance Type 1 and Type 2 errors
- Interpret p-values
- Use p-values in one-tailed and two-tailed hypothesis testing

Significance Levels

The significance level α is also the probability of incorrectly rejecting the null hypothesis $H_0 = \mu_0$ if the null is true.

Let us assume that we are willing to accept a 5% probability of incorrectly rejecting the null (α = 0.05).

Then z_α= 1.645.

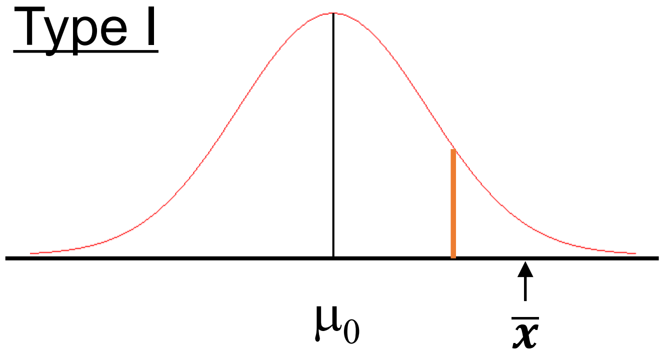


Type I and Type II Errors

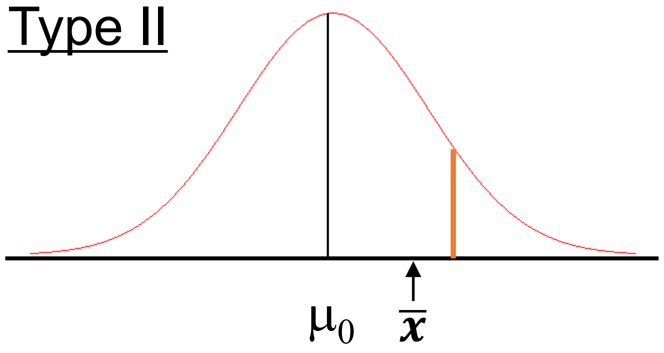
Key idea: Making inferences about the population parameters based on sample statistics is inherently uncertain and thus subject to error.

Our decision

	Do Not Reject H_0	Reject H_0
Reality		



H_0 is True



H_0 is False

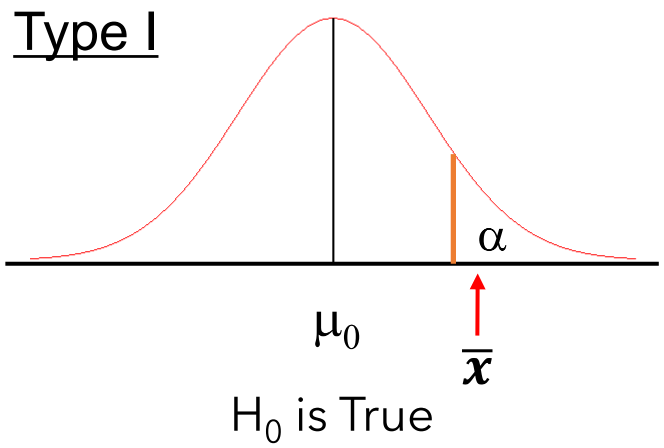
Probabilities of Type I and Type II Errors

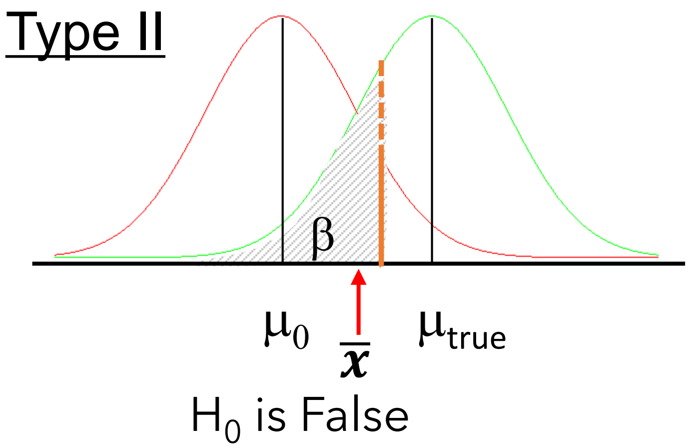
Key idea: Making inferences about the population parameters based on sample statistics is inherently uncertain and thus subject to error.

Let α = probability of making a type I error (rejecting a true null)
Let β = probability of making a type II error (failing to reject a false null)

As discussed previously, α is the total probability in the tails of the null distribution.

β is hard to calculate: it depends on how far the true mean (μ) is from the null mean (μ_0).





Type I and Type II Errors

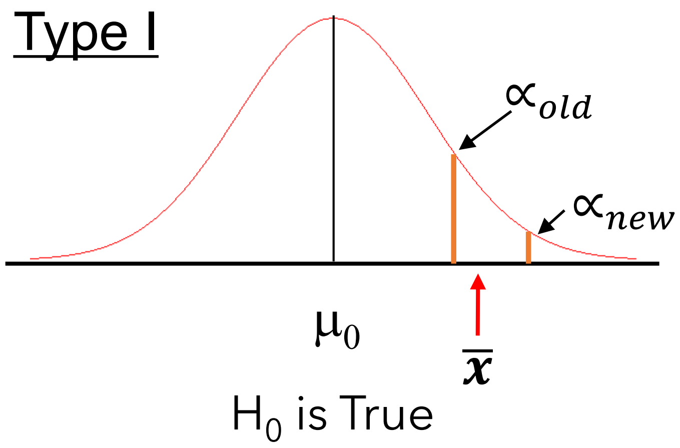
Key idea: Making inferences about the population parameters based on sample statistics is inherently uncertain and thus subject to error.

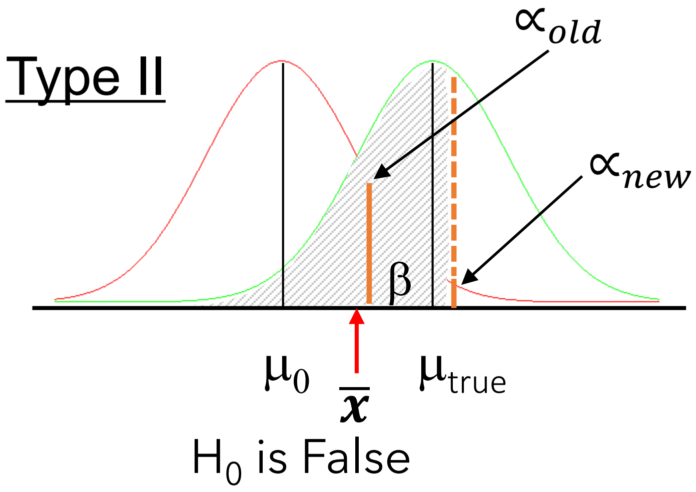
Let α = probability of making a type I error (rejecting a true null)
Let β = probability of making a type II error (failing to reject a false null)

As discussed previously, α is the total probability in the tails of the null distribution.

β is hard to calculate: it depends on how far the true mean (μ) is from the null mean (μ_0).

For a fixed sample size decreasing α necessarily increases β. Thus, there is a trade-off between Type I and II errors.





Tests of Understanding on the Consequences of Making Type I & II Errors

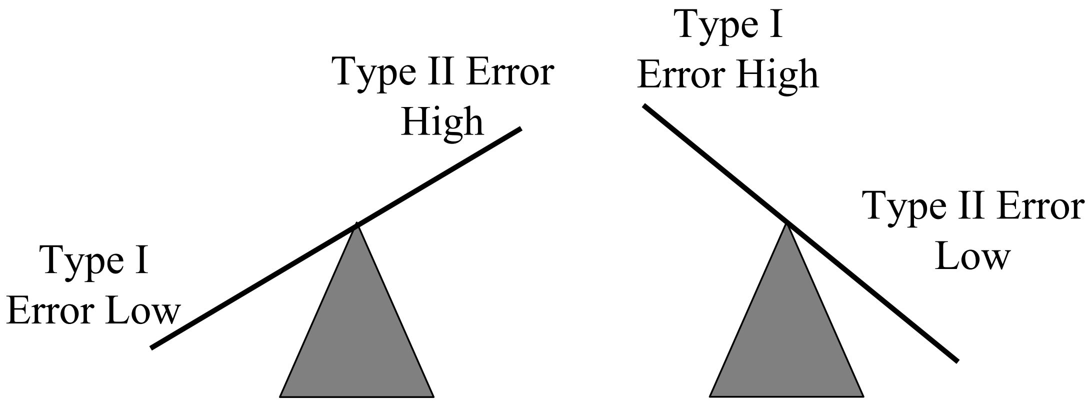
- A computer supplies retail chain has a policy of only opening stores in communities where households spend more than $40 per year on computing supplies and equipment.

$$H_0 : \mu \leq 40$$

$$H_1 : \mu > 40$$

- In the context of this problem, what is a Type I error, and what are its consequences?
- In the context of this problem, what is a Type II error, and what are its consequences?
- How can we reduce the probability of making a Type II error without increasing the probability of a Type I error?

Balancing Type I & Type II Errors



- List and, to the extent possible, quantify the costs of a Type I error.
- List and, to the extent possible, quantify the costs of a Type II error.
- Choose a value of α that reasonably balances these costs. This may also require hiring an expert to calculate β and also collecting more data which may be costly.

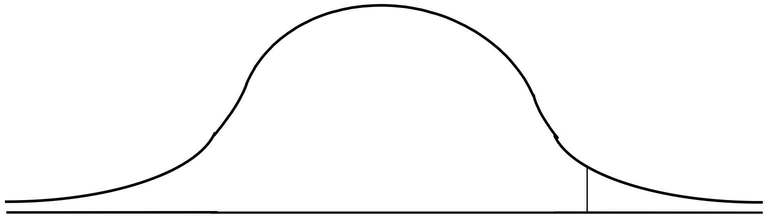
P-Values

- Provide a quantitative measure of the strength of evidence in support of the conclusion of a hypothesis test
- If we reject null: Was it a close call or firmly in favor of the alternative?
- If we fail to reject null: How close did we come to rejecting?

P-Values for Tests of μ

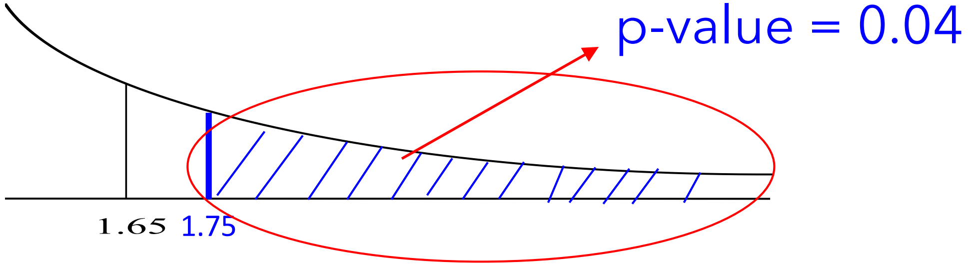
The P-value is the smallest level of significance for which the observed sample statistic tells us to reject H_0

Example:
$H_0: \mu \leq 0$
$H_a: \mu > 0$
If $\alpha = 0.05$



Sample evidence: $\bar{x} = 3.5$ s = 30
n = 225
t = 1.75

What is $P(t \geq 1.75)$?   0.04

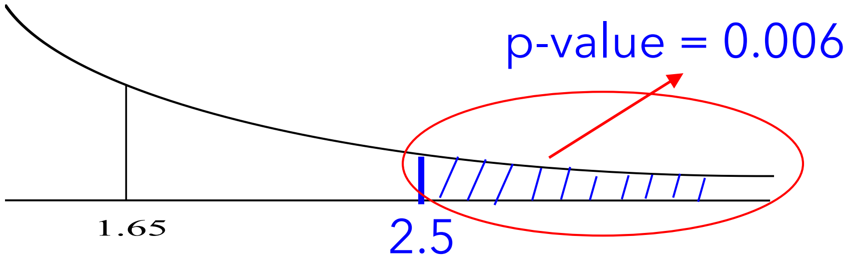


<u>Important Point:</u>
If P-value ≤ α, we reject null hypothesis

P-Values Continued

Suppose instead: $\bar{x}$ = 5.0 s = 30 n = 225 → t = 2.5

What is P(t ≥ 2.5)? 0.006

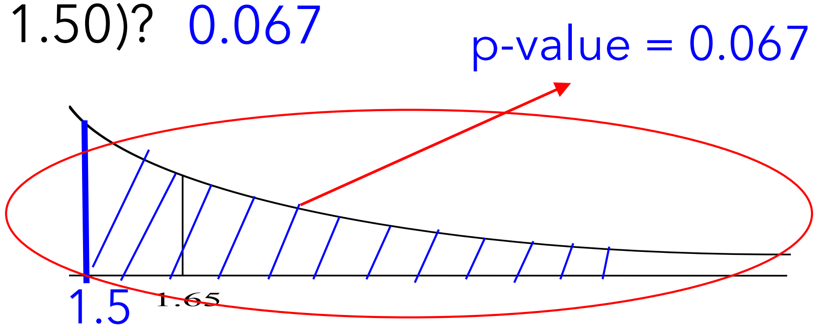


<u>Conclusion</u>

Reject because the p-value is less than α

Now suppose: $\bar{x}$ = 3.0 s = 30 n = 225 → t = 1.5

What is P(t ≥ 1.50)? 0.067



<u>Conclusion</u>

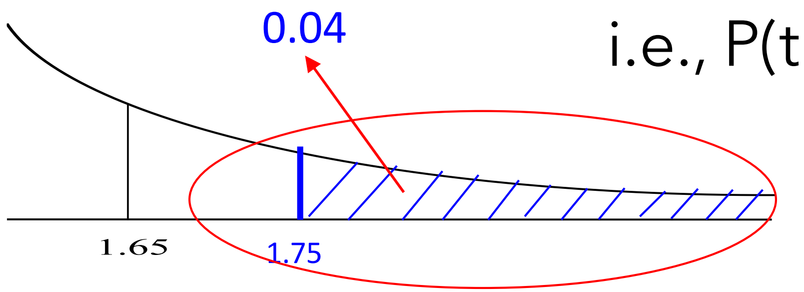
Do not reject because the p-value is more than α

P-Values and α

If p-value **≤** α we reject null, the alternative is "statistically significant."

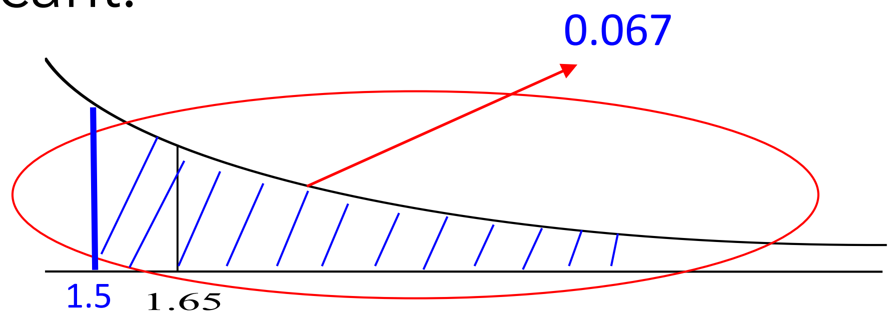
For $\alpha = 0.05$, $t_{0.05} \cong 1.65$
i.e., $P(t \geq 1.65) = 0.05$

$t = 1.75$
$P(t \geq 1.75) = 0.04$



If p-value > α we don't reject null—alternative is said "not to be statistically significant."

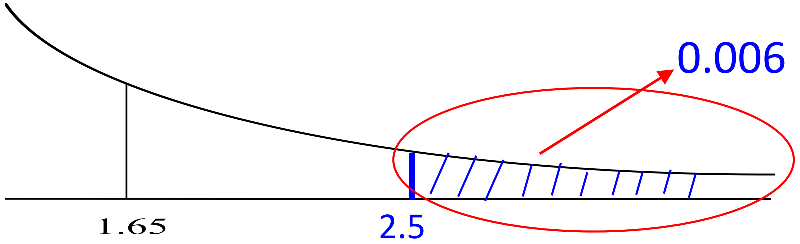
$z = 1.50$
$P(t \geq 1.50) = 0.067$



Interpreting P-Values

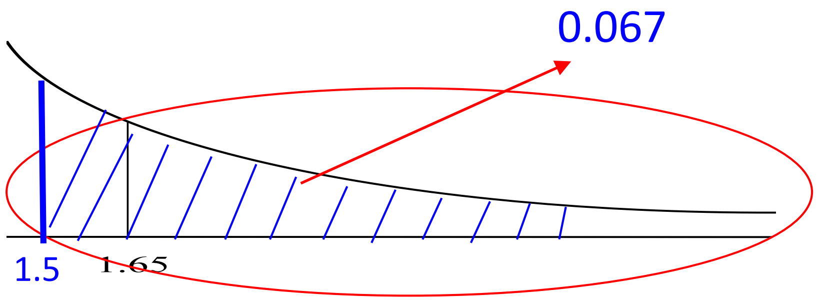
If p-value is a lot smaller than α, alternative is commonly described as “highly significant.”

t = 2.50
P(t ≥ 2.50) = 0.006



If p-value is only a little bit bigger than α , alternative is commonly said to be “almost significant.”

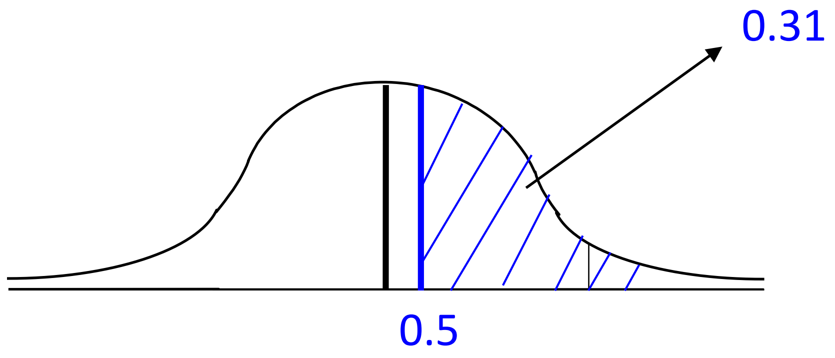
t = 1.50
P(t ≥ 1.50) = 0.067



Interpreting P-Values Continued

If p-value is a lot bigger than α, alternative is commonly said to be "not even close to being significant."

$t = 0.50$
$P(t \geq 0.50) = 0.31$



P-Values for 2-Tailed Tests

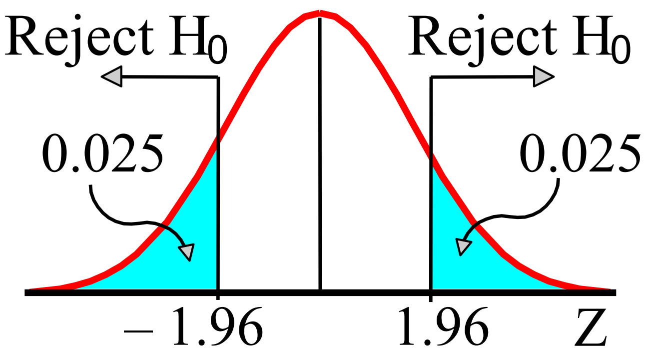
Step 1: Determine Hypotheses and critical values.

H_0: $\mu=100$

H_1: $\mu\neq100$

$\alpha = 0.05$ $\bar{x} = 103.5$

$n = 400$ $s = 33.3$



Step 2: Calculate the test statistic

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = 2.10$$

Step 3: Calculate p-value

P(t≥ 2.10) = 0.018

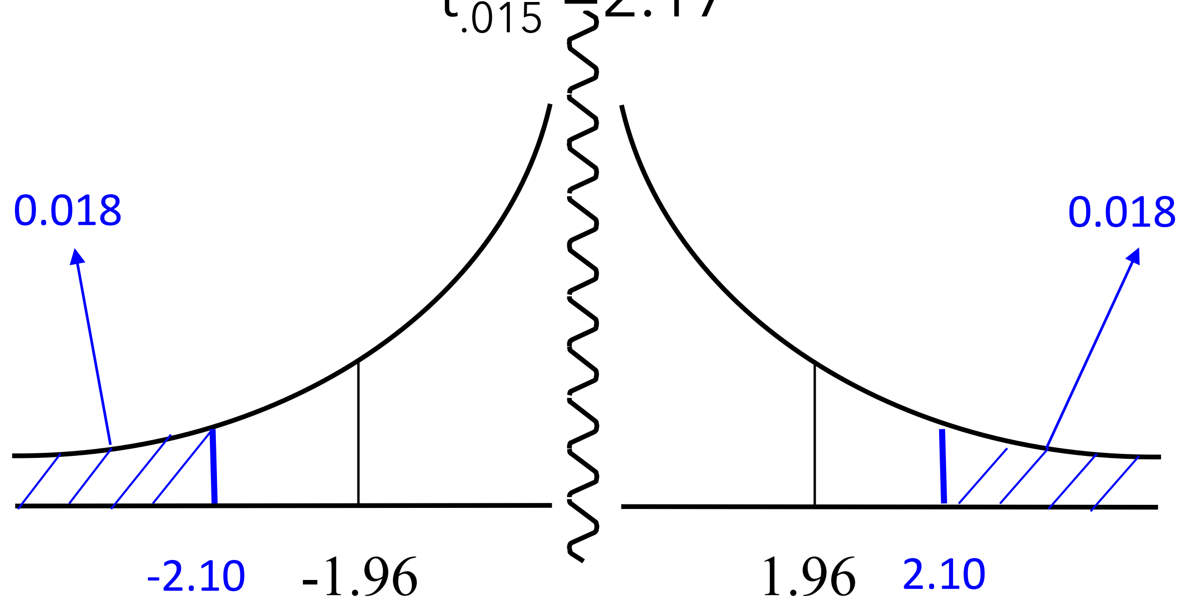
Step 4: Draw conclusion

Reject Null

Two-Tail P-Value

What's the conclusion of the hypothesis test for $\alpha = 0.03$?

$t_{.015} = 2.17$



Important!!

When calculating a p-value for a two-tail test, **ALWAYS** double the value of the p-value before comparing to α.

P-Values for Tests of Population Proportions

- Calculated with z-scores, not t-scores
- Same Interpretation as P-values about test of μ

P-Values in ISLE

No Filters

Hypothesis test for calinfo:

Let p be the population probability of calinfo being 0.

We test

$$H_0 : p = 0.5 \; vs. \; H_1 : p \neq 0.5$$

Sample proportion: 0.48

```
One-sample z-test

Alternative hypothesis: True mean is not equal to 0.5

    pValue: 0.3033
    statistic: -1.0293
    95% confidence interval: [0.4408,0.5184]

Test Decision: Fail to reject null in favor of alternative at 5% sign
ificance level
```

Hypothesis test for AGE:

$$H_0 : \mu = 27 \; vs. \; H_1 : \mu \neq 27$$

```
One-sample t-test

Alternative hypothesis: True mean is not equal to 27

    pValue: 0.0003
    statistic: 3.608
    df: 634
    95% confidence interval: [27.8239,29.7918]

Test Decision: Reject null in favor of alternative at 5% significance level
```

Tests of Understanding for P-Value

- Suppose the p-value of a one-tailed test of μ is .015. What is the minimum value of α for which the null hypothesis can be rejected? What is the corresponding t score cut-off for this test if n=144?

- Suppose the p-value of a two-tailed test of a population proportion is .015. What is the minimum value of α for which the null hypothesis can be rejected? What is the corresponding z score cut-off for this test?