Data Analysis 4 Hint:

On part 2, question a, you are asked to find the 10th and 90th percentiles of a Standard Normal Distribution. The Standard Normal is a special case of the normal distribution with a fixed mean and standard deviation. We discussed this distribution initially on Tuesday of week 4 in the Normal Distribution notes.

You will need your answer from part a to answer part b.

DA 4 is due tonight!

Week 6

Hypothesis Testing

ST 314
Introduction to Statistics for Engineers

Oregon State University

Estimation vs. Hypothesis Testing

Want to know more about a population characteristic. → Goal is to estimate the population characteristic (unknown parameter). → **Estimation** Point estimates and confidence intervals

Have a claim or guess about a population characteristic. → Goal is to use sampled data to test the validity of the claim. → **Hypothesis Testing** Hypotheses, test statistics, and p-values

Examples of when to use a hypothesis test:

- Test the average train ride time against the advertised time
- Test the proportion of defective parts in a manufacturing process against a specified standard
- Compare whether the average tensile strength of rubber seals is different between machines

Hypothesis Test Example

A high-speed train between Florence and Rome, Italy is advertised to take 94 minutes.

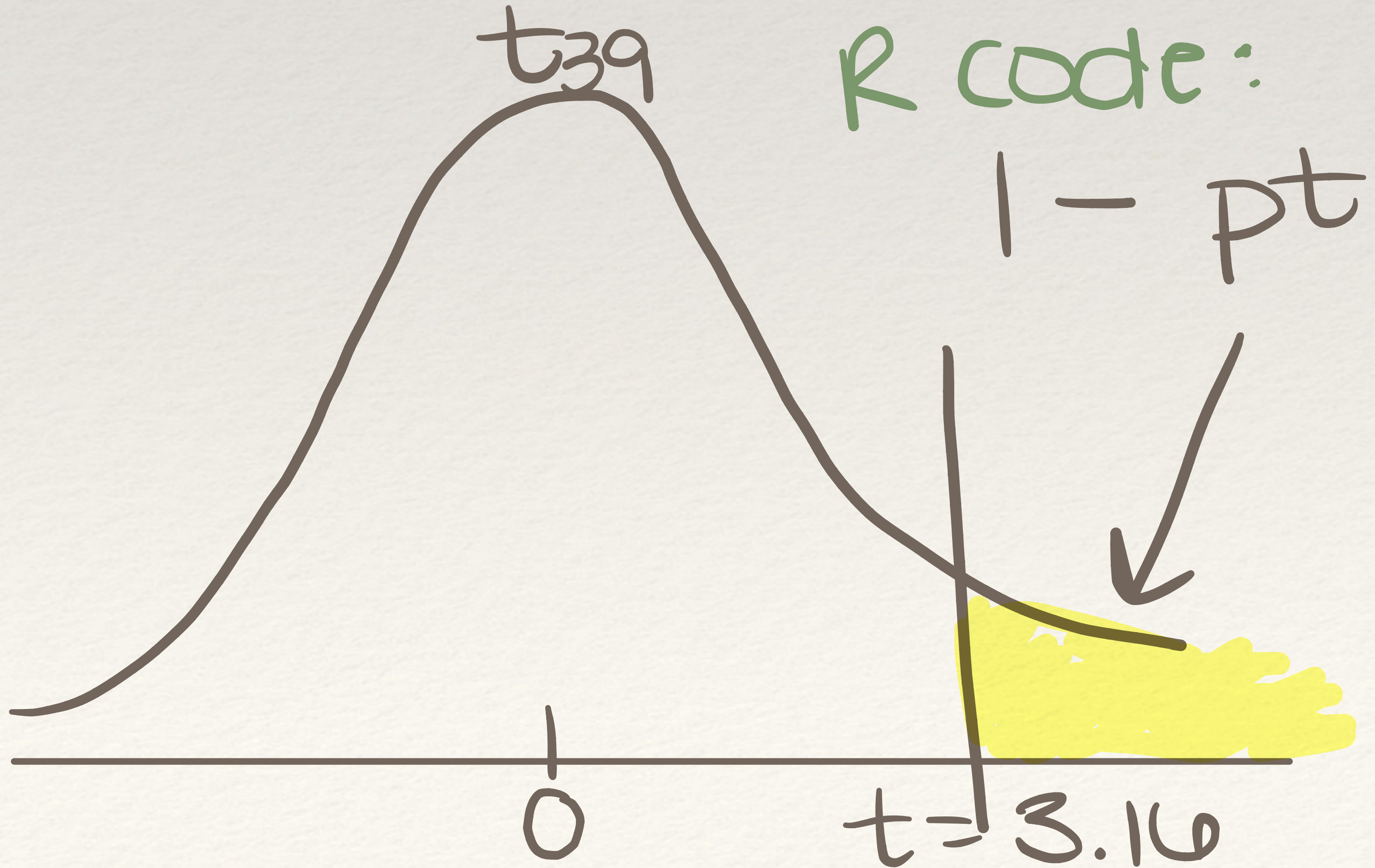
A frequent rider is consistently late and is convinced the average ride time is longer. She decided to record the time of 40 random rides. Her sampled average is 99 minutes. The sample has a standard deviation of 10 minutes.

$\bar{x} = 99 \quad n = 40 \quad df = 39$

$s = 10 \quad \mu = 94$

If the advertised time of 94 minutes is the actual average, what is the chance the mean from a sample of 40 rides would exceed 99 minutes?

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{99 - 94}{10/\sqrt{40}} = 3.16 \sim t_{39}$$



R code:

1 - pt(3.16, 39)

= 0.0015

Steps in Performing a Hypothesis Test

1. State question of interest
2. Identify the parameter (μ or p)
3. State the null and alternative hypotheses
4. Check the conditions for the test
5. Determine the null distribution
6. Calculate test statistic + p-value
7. Make a conclusion

The Hypotheses

The Null Hypothesis

- Establishes a claim
- Assume the hypothesis true when performing the test
- Always a statement of equality (=)
- denoted H_0

Ex. $H_0: \mu = 94$

The Alternative Hypothesis

- Reflects the question of interest
- Contradicts the null
- is either one- or two-sided
- denoted H_A (or H_1)

Ex. $H_A: \mu > 94$

Always in terms of the parameter!

Hypothesis Test Example

A high-speed train between Florence and Rome, Italy is advertised to take 94 minutes.

A frequent rider is consistently late and is convinced the average ride time is longer. She decided to record the time of 40 random rides. Her sampled average is 99 minutes. The sample has a standard deviation of 10 minutes.

Determine the null and alternative hypotheses for this test based on the train site's claim that rides take 94 minutes.

$H_0: \mu = 94$

$H_A: \mu > 94$

μ = average ride time

One- and Two-sided Alternatives

μ_0 = claimed mean

p_0 = claimed proportion

Lower one-sided alternative

Question of interest: is the parameter **less than** the claimed value?

$H_A: \mu < \mu_0$ | $H_A: p < p_0$

Upper one-sided alternative

Question of interest: is the parameter **greater than** the claimed value?

$H_A: \mu > \mu_0$ | $H_A: p > p_0$

Two-sided alternative

Question of interest: does the parameter **differ** from the claimed value?

$H_A: \mu \neq \mu_0$ | $H_A: p \neq p_0$

One- or Two-sided Practice

Consider the following alternative hypothesis examples. Which are one-sided? Which are two-sided?

1. The average time to finish a maze is not equal to 16 minutes. Two-sided
2. The average running pace is less than 7 mph. Lower one-sided
3. The average IQ score of OSU students is greater than 119. Upper one-sided
4. The average exam score is different than 75%. Two-sided
5. The average lunch break is less than 30 minutes. Lower one-sided

Test Statistics ~~& p-values~~

Similar to "standardizing", a **test statistic** compares a sample statistic to a hypothesized value while accounting for the variability of the statistic.

The general form of a test statistic is:

$$\frac{\text{estimate} - \text{claim}}{\text{standard error}}$$

The Null Distribution

assume to be true

- $H_0 : \mu = \mu_0$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

sample proportion

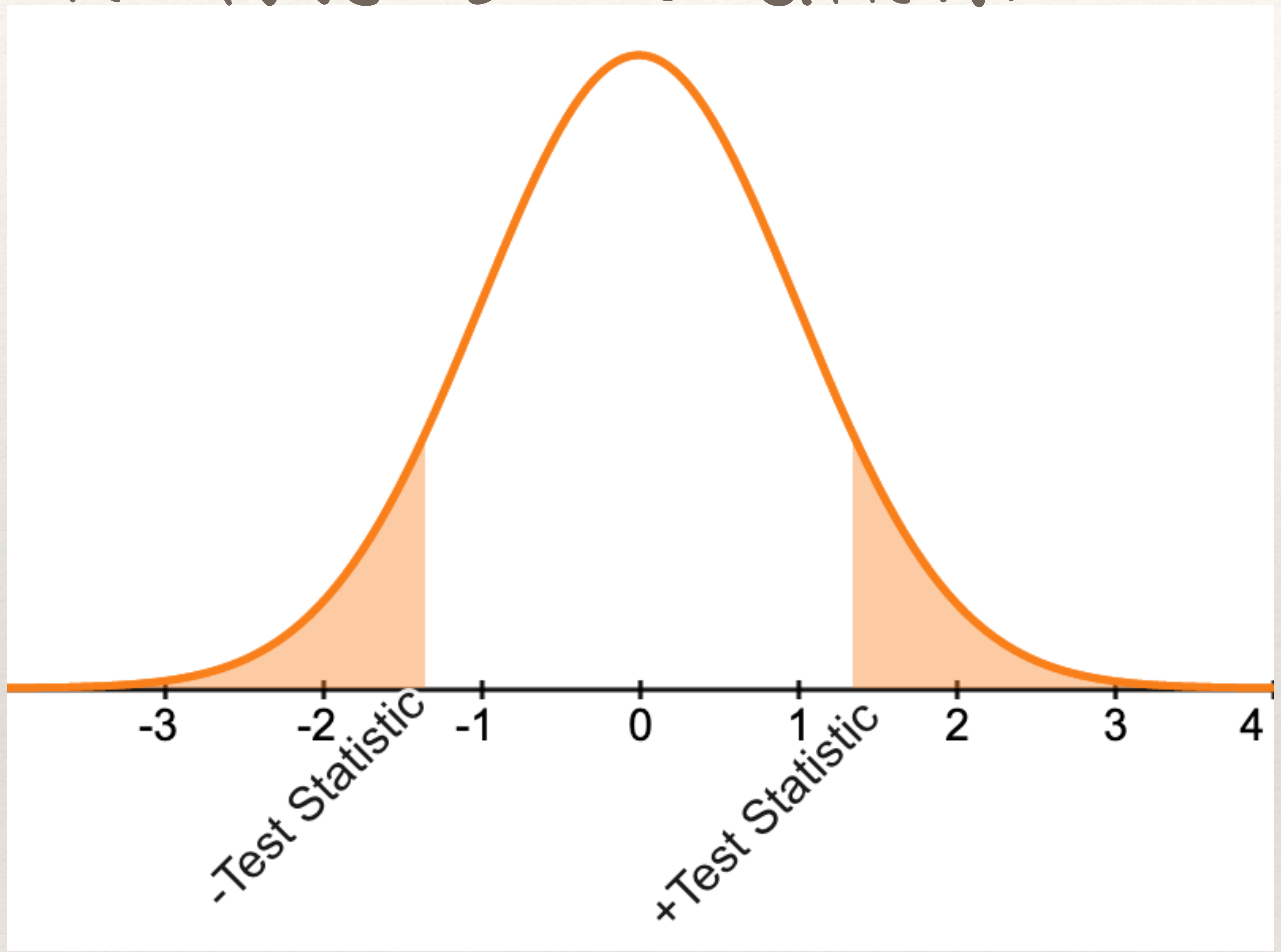
claimed value

- $H_0 : p = p_0$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0, 1)$$

p-values

Ex. Two-sided alternative



- The p-value is a probability. It is the area under the tail(s) of the null distribution curve.

Assuming the hypothesized parameter value is true, the probability the sample statistic would take on a value as or more extreme than the one we actually observed is the p-value.

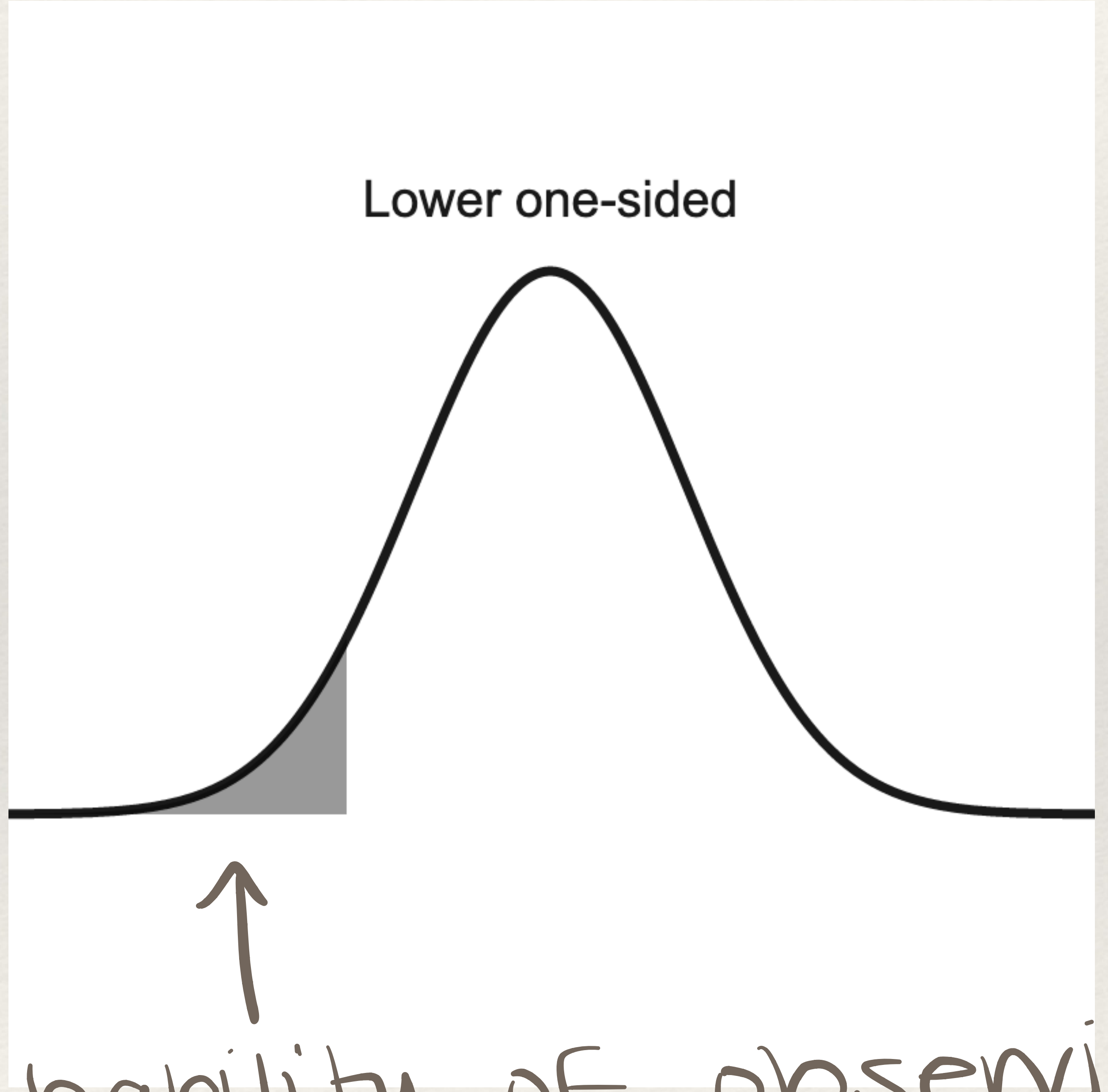
Ex. $\underbrace{P(\bar{X} \geq 99) \text{ when } \mu = 94}$

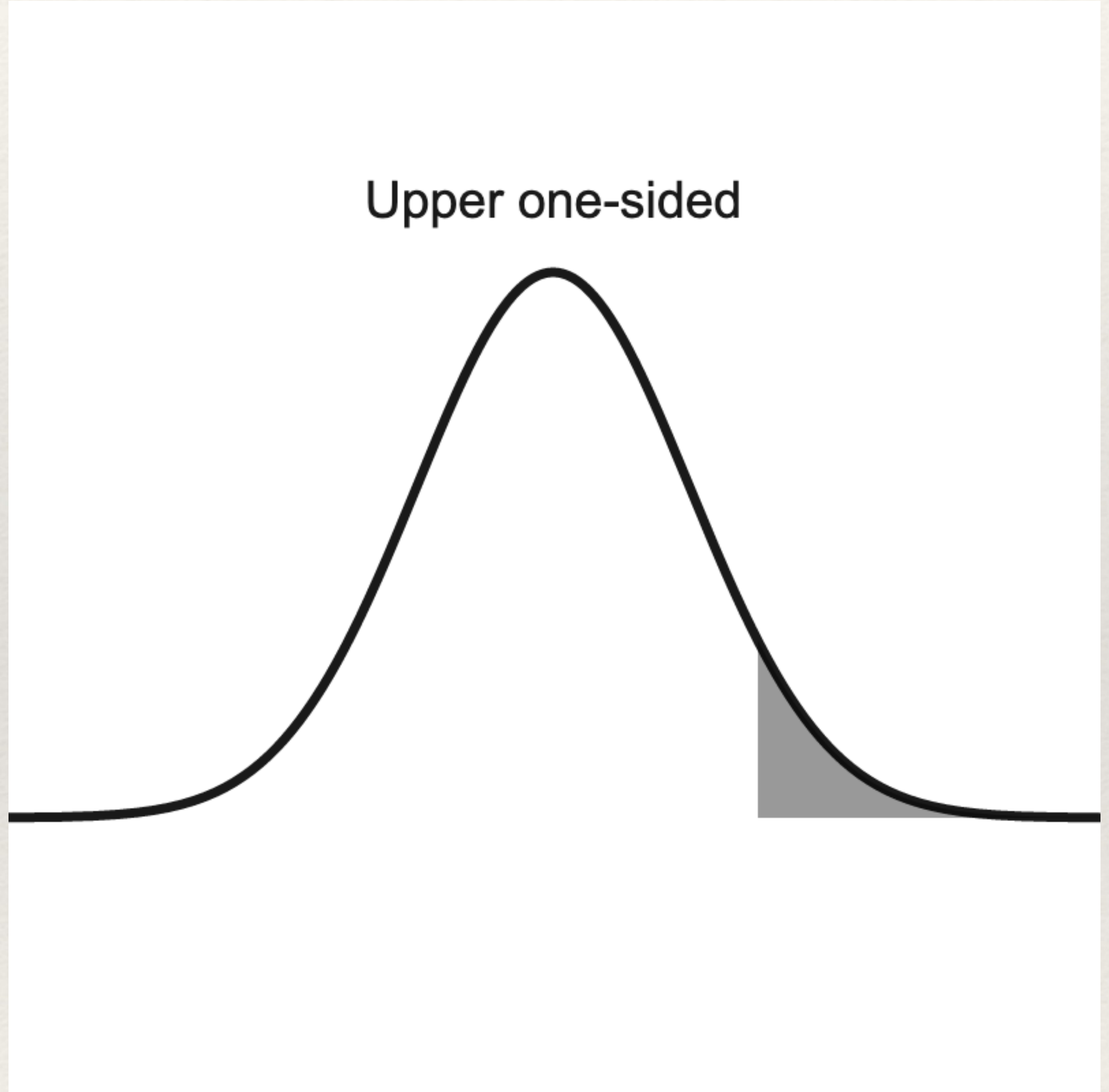
p-value = 0.0015

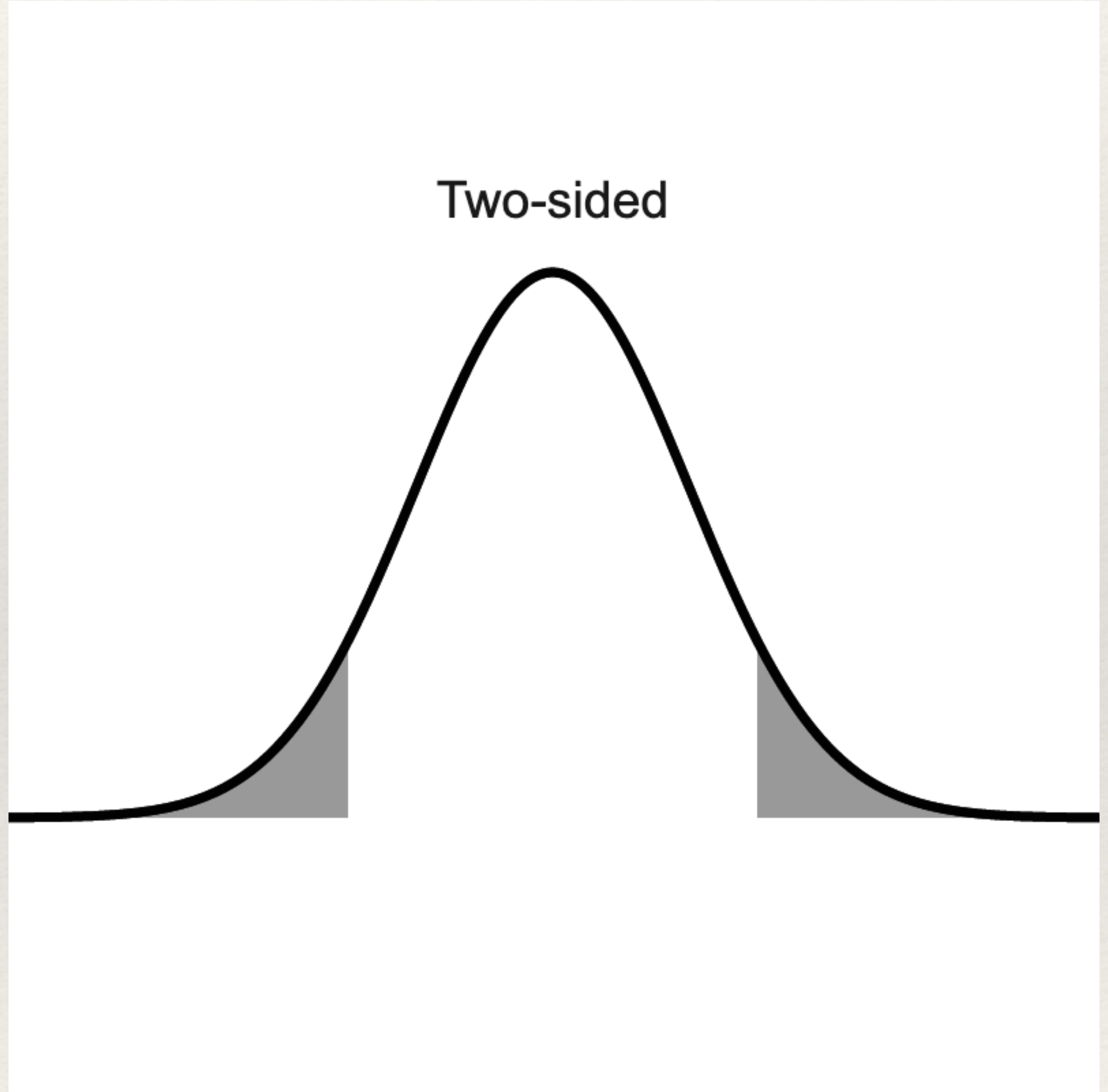
One- and Two-sided p-values

$H_A: \mu < \mu_0$

$H_A: \mu > \mu_0$







probability of observing a sample mean less than or equal to $\bar{x}$ when the true pop. mean is μ_0

t-Test for a Mean

- When to use: test average value of μ for a single population
- Conditions required for inference:
  - sample is representative of population (random)
  - n sufficiently large (central limit theorem)
- Null & Alternative hypotheses: $H_0: \mu = \mu_0$ ← replace this with hypothesized value

$H_A: \mu < \mu_0$ OR $H_A: \mu > \mu_0$ OR $H_A: \mu \neq \mu_0$

- Test statistic: $t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$

$H_0: \mu = 94$

$H_a: \mu > 94$

t-Test for a Mean Example

A high-speed train between Florence and Rome, Italy is advertised to take 94 minutes.

A frequent rider is consistently late and is convinced the average ride time is longer. She decided to record the time of 40 random rides. Her sampled average is 99 minutes. The sample has a standard deviation of 10 minutes.

$n = 40$ ✓

We want to perform a hypothesis test to test whether the average ride time is greater than 94 minutes. To do so, first check the conditions. If they are met, proceed to calculating the test statistic.

$$t = \frac{99 - 94}{10/\sqrt{40}} = 3.16$$

$H_0: \mu = 94$

$H_A: \mu > 94$

t-Test for a Mean Example

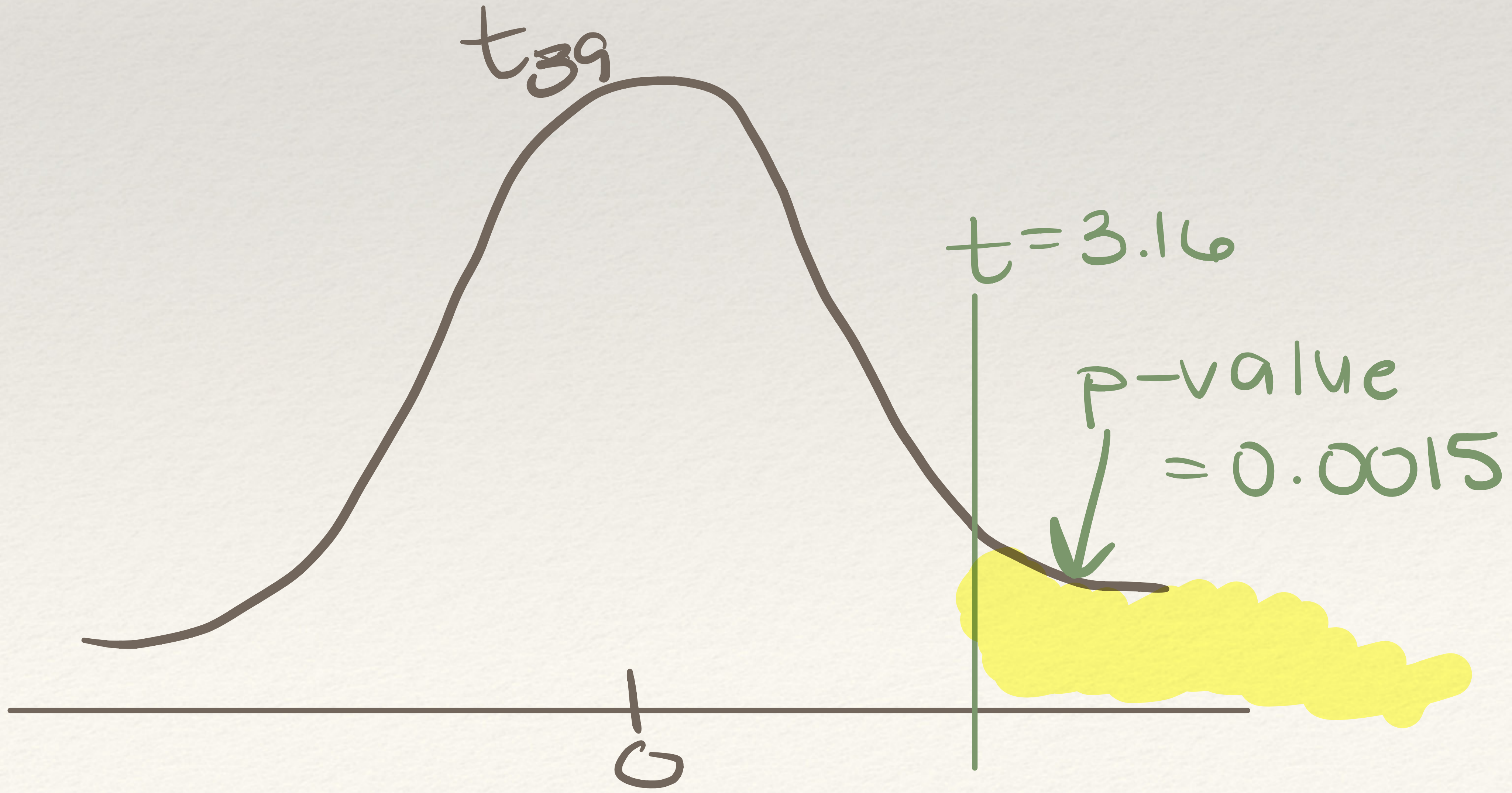
A high-speed train between Florence and Rome, Italy is advertised to take 94 minutes.

A frequent rider is consistently late and is convinced the average ride time is longer. She decided to record the time of 40 random rides. Her sampled average is 99 minutes. The sample has a standard deviation of 10 minutes.

Now that we've calculated the test statistic, we know how many standard deviations away from the hypothesized mean our sample mean is. We can use this value to calculate the probability we would have observed a sample mean as or more extreme than 99 minutes if we assume the true population mean to be 94 minutes (this calculation is the p-value). Calculate the p-value.

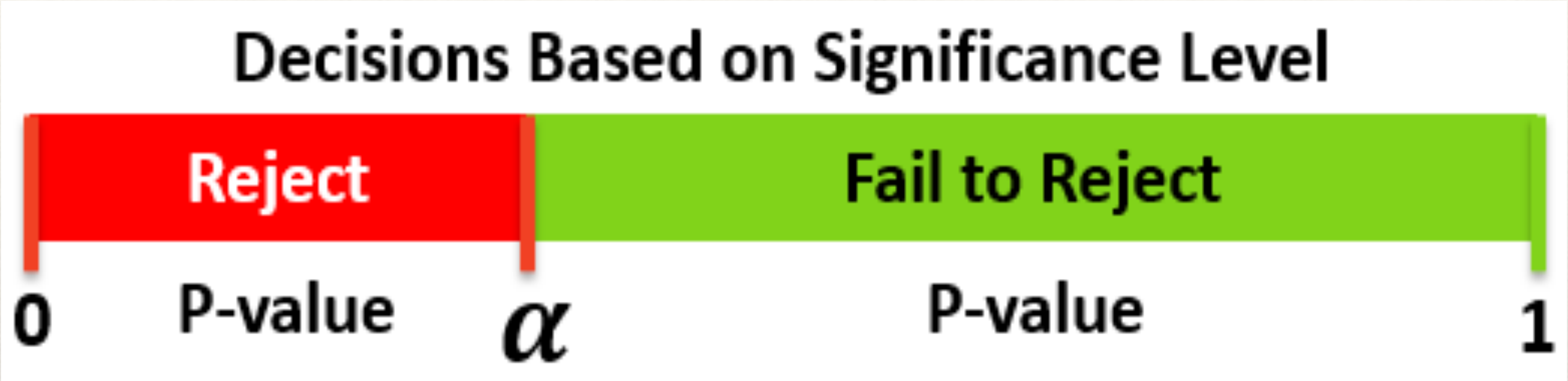
$t = 3.16$

$df = 40 - 1$
$= 39$



Decisions Based on p-values

The **significance level**, denoted by α, is a predetermined cut off to which we will reject the null hypothesis. Common values for α are:

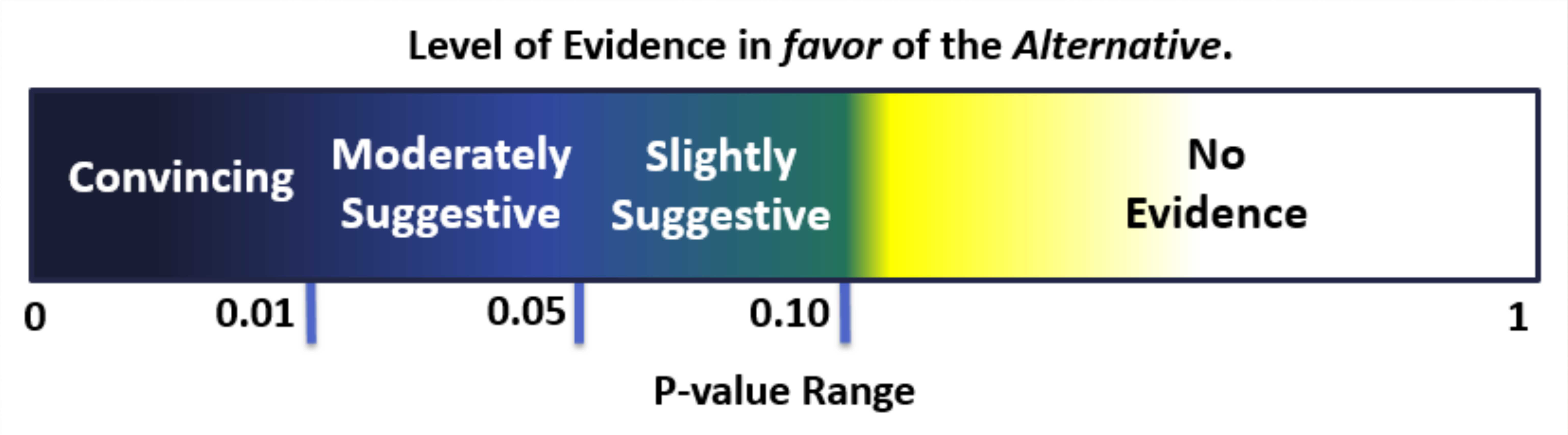


A ________________ α requires a more extreme test statistic to reject the null. Decreasing α will reject less often, but will decrease the chance of error.

Never say the null is "false". Instead, we say "the evidence is very strong to reject the null"

Statements in terms of the alternative

- Using only terms like “reject” and “fail to reject” in conclusions may confuse novice readers.
- We’ll provide a more complete conclusion by providing a statement of evidence in terms of the alternative hypothesis that reflects the question of interest.



Four-part Conclusion

A good conclusion must have ______________ and should include the following four important pieces.

1. Statement for the ____ for the alternative hypothesis.
2. Whether to ________________ or ________________________________ the null hypothesis based on α.
3. The __________________ estimate for the parameter of interest.
4. A $(1-\alpha)100\,\%$ ____________________________________ estimate for the parameter of interest.

Hypothesis Test Conclusion Example

A high-speed train between Florence and Rome, Italy is advertised to take 94 minutes.

A frequent rider is consistently late and is convinced the average ride time is longer. She decided to record the time of 40 random rides. Her sampled average is 99 minutes. The sample has a standard deviation of 10 minutes.

Write a four-part conclusion for this hypothesis test.

Errors in Hypothesis Testing

	H_0 True	H_0 False
Reject H_0	Type I Error	Correct Conclusion
Fail to reject H_0	Correct Conclusion	Type II Error

- When a test performs as it should, the value α is the probability of making a ________________ error.
- When the consequences are high, reduce the chance of making a Type I error by choosing a ________________ significance level.

Type I and Type II are related. Decreasing the chance of one error will increase the chance of the other.

z-Test for a Proportion

- When to use:
- Conditions required for inference:
- Null & Alternative hypotheses:
- Test statistic:

z-Test Example

Pew Research asked a random sample of 1000 American adults whether they supported the increased usage of coal to produce energy. Set up hypotheses to evaluate whether a majority of American adults support or oppose the increased usage of coal.

z-Test Example

Confidence Intervals vs. Hypothesis Tests

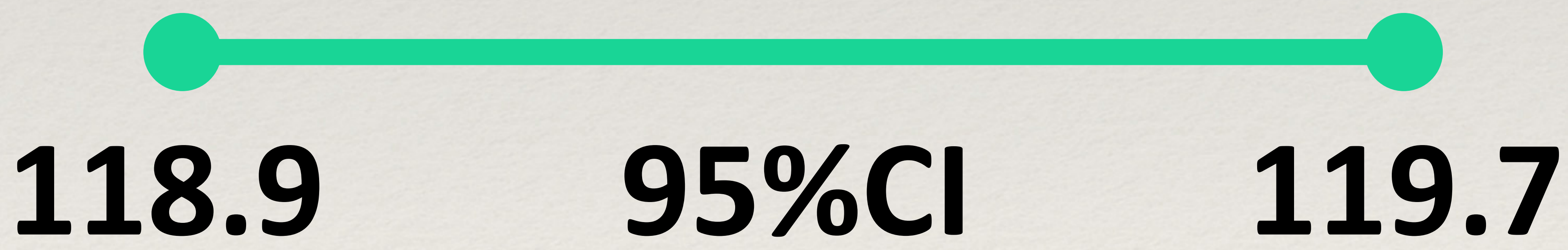
If a $(1-\alpha)100\,\%$ confidence interval, does not contain μ_0, the hypothesized value, then μ_0 it is not a ____________________ for our estimate. The two-sided hypothesis test will ____________________ the null at the α significance level.

Example

Suppose a manufacturer of boards for construction would like to test the length of a batch of boards with the following null and alternative hypotheses:

$H_0 : \mu = 120$ inches
$H_A : \mu \neq 120$ inches

$$\mu_0 = 120$$



Reject at 5% Significance!