Overview

- Null and Alternate Hypotheses
- Hypothesis Tests, one-tail and two-tailed
- p-value
- Significance Level

Writing the Hypotheses

Example 1. Write the null and alternative hypotheses for the following situations in symbols and in words.

a. A current antacid provides relief for 70% of the people who use it. A pharmaceutical company has a new drug and they want to test whether it is more effective.

$H_0: p = .70$ The new drug is the same as the old drug (no more effective)

$H_A: p > .70$ The new drug is more effective more than 70% get relief.

The Null Hypothesis, H_0

The null hypothesis is the current value of the parameter, the accepted value, or the status quo. The form is $\boldsymbol{H_0: p = value}$

The Alternate Hypothesis, H_A or H_1

The alternate hypothesis is what we are trying to prove. The possible forms of the alternate hypothesis are $\boldsymbol{H_A: p > value}$

$\boldsymbol{H_A: p < value}$

$\boldsymbol{H_A: p \neq value}$

b. A company is concerned that too few of its cars meet pollution standards. They want to test whether less than 80% of their fleet meets emissions standards.

$H_0: p = .80$

$H_A: p < .80$ ← the thing they want to test or prove

c. In 2014, the official poverty rate in the U.S. was 14.8% according to the U.S. Census Bureau. A local official wants to test whether their county has a different poverty rate than the rest of the U.S.

$H_0: p = .148$

$H_A: p \neq .148$

Testing the Hypothesis

Example 2. Let's continue the antacid example. A current antacid provides relief for 70% of the people who use it. A pharmaceutical company has a new drug and they want to test whether it is more effective. They run a study with 100 randomly selected patients and 75 people experienced relief. Is this evidence that the new drug is better?

Step 1. Write the Hypotheses

$H_0: p = .70$

$H_A: p > .70$

Step 2. Check the four conditions needed to use a sampling distribution (Normal)

1. independence – patients would be independent on whether the drug was effective
2. randomization – random sample
3. 10% condition – 100 people is less than 10% of the population who might use this drug
4. success/failure – $np = 100(.70) = 70 \checkmark \geq 10$, $nq = 100(.30) = 30 \geq 10 \checkmark$

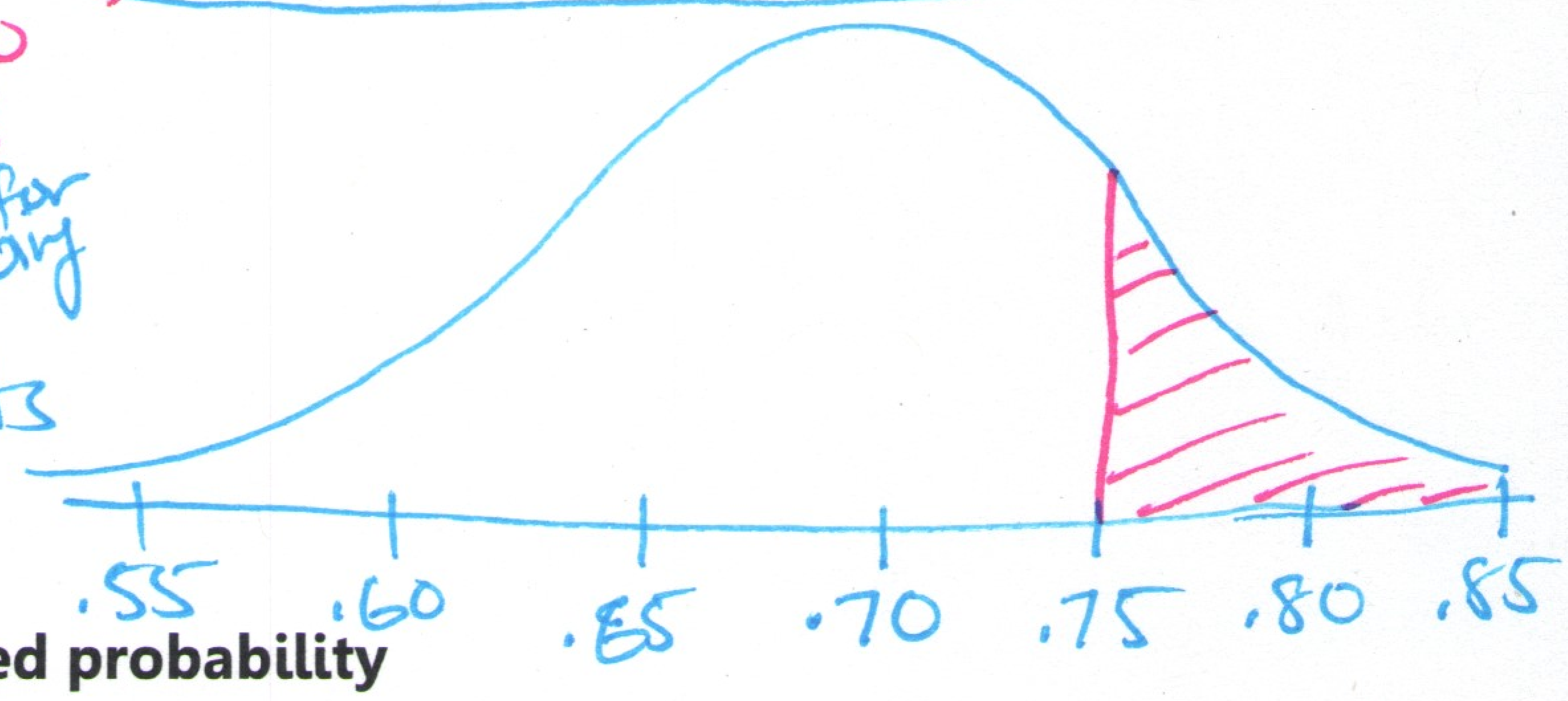
Step 3. Draw the Normal model, assuming H_0 is true

$\hat{p} \sim N\left(p, \sqrt{\frac{pq}{n}}\right)$ $\hat{p} \sim N\left(.70, \sqrt{\frac{(.70)(.30)}{100}}\right)$

.0458

.05 for drawing

proportion of patients experiencing relief in samples of 100



Step 4. Find $\hat{p}$ and shade the one-tailed or two-tailed probability

$\hat{p} = \frac{75}{100} = .75$

Shade to the right because H_A is greater than.
one-tailed test (>)

Step 5. Calculate the p-value and compare with the significance level $\alpha = .05$ unless stated otherwise 5% chance

p-value $= P(\hat{p} \geq .75) = .1375$

The probability of getting the observed result, assuming the null hypothesis is true. $.1375 \not< .05$

Step 6. State a clear conclusion that includes the p-value and whether you reject the null or fail to reject the null and why

not statistically significant

The p-value is .1375, which is not less than .05. The result is not statistically significant. We fail to reject the null hypothesis. We do not have enough evidence to suggest the new drug is better.

Steps to a Hypothesis Test

1. Write the null and alternate hypotheses
2. Check the 4 conditions required to use the Sampling Distribution for $\hat{p}$ (Normal)
3. Draw the Normal Model assuming H_0 is true
4. Find $\hat{p}$ and shade the one-tailed or two-tailed probability
5. Calculate the p-value and compare with the significance level, α
6. State a clear conclusion that includes the p-value and whether you reject the null or fail to reject it and why

What is a p-value?

The p-value is the probability of getting the observed result, $\hat{p}$, given that the null is true. If the chance of getting that result is small enough then we have enough evidence to reject the null hypothesis

Significance Level, alpha or α

The most common significance level is 0.05. If there is less than a 5% chance of getting the observed result, then that is statistically significant. For medical tests we may use $\alpha = 0.01$. In the social sciences we may use use $\alpha = 0.10$. Unless otherwise stated, use $\alpha = 0.05$.

Stating your Conclusion ✱

If the **p-value ≥ α,** we fail to reject the null hypothesis. There is not enough evidence to suggest that the alternate hypothesis is true. (It is likely that the result is due to random variation, so the difference is not statistically significant.)

If the **p-value < α,** we reject the null hypothesis. There is evidence to suggest that the alternate hypothesis is true. (It is unlikely that the result is due to random variation alone, so it is statistically significant.)

Why do we Fail to Reject?

You might be wondering why we don't "accept" the null. We can only fail to reject it. Think about an example using the court system. A defendant is innocent until proven guilty. The burden of proof lies with the alternative hypothesis.

H_0: The defendant is innocent
H_A: The defendant is guilty

If there is sufficient evidence than the defendant may be proven guilty (reject the null hypothesis). Otherwise, we fail to reject the null, and they are proven not guilty. It is impossible to prove that someone is innocent, even though they very well may be.

Statistical Significance vs. Practical Significance

If a very large sample size is used, then very small differences can be statistically significant. The difference may not be meaningful. In later courses, you may learn how to choose the sample size so that the statistical significance reflects a meaningful or practical difference.

Comparison with a Confidence Interval

Instead of doing a hypothesis test we could take the value of $\hat{p}$ and form a 95% confidence interval. This would be similar to doing a two-tailed test at a 5% significance level. The main difference is that in a confidence interval the standard deviation is calculated with $\hat{p}$, but in hypothesis testing the standard deviation is calculated using the value from the null hypothesis, p_0. If p_0 is close to $\hat{p}$ then they are very close.

Example 3. Let's continue the emissions example. A company is concerned that too few of its cars meet pollution standards. They want to test whether less than 80% of their fleet meets emissions standards. They measure a random sample of 500 vehicles and 385 meet the standards. Is this evidence that too few of its cars meet the standards?

Step 1. Write the Hypotheses

$H_0: p = .80$

$H_A: p < .80$

Step 2. Check the four conditions needed to use a sampling distribution (Normal)

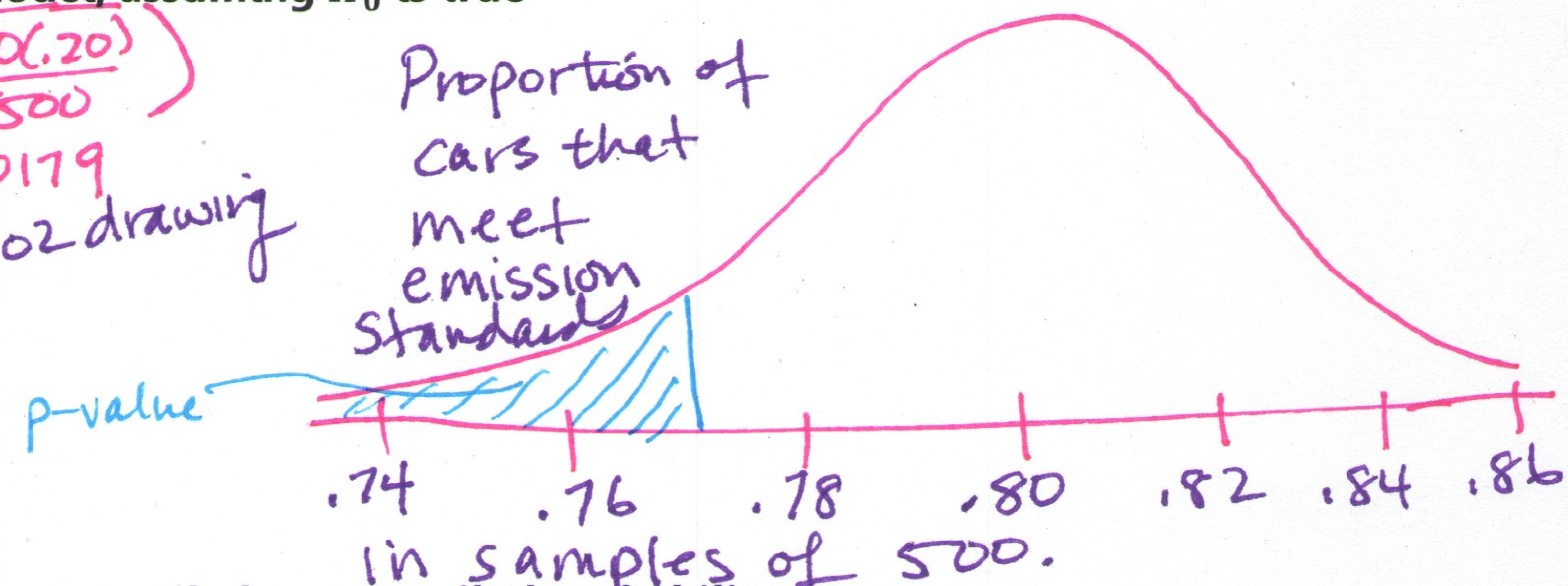
① independence – yes if they choose different models (stratify)
② randomization – random sample
③ <10% – 500 is less than 10% of all cars the company makes
④ $np = 500(.80) = 400 \geq 10$ ✓ $nq = 500(.20) = 100 \geq 10$ ✓

Step 3. Draw the Normal model, assuming H_0 is true

$\hat{p} \sim N\left(.80, \sqrt{\frac{.80(.20)}{500}}\right)$

.0179

.02 drawing



Step 4. Find $\hat{p}$ and shade the one-tailed or two-tailed probability

$\hat{p} = \frac{385}{500} = .77$ one-tailed test

Step 5. Calculate the p-value and compare with the significance level

$P(\hat{p} \leq .77) = .0469 < .05$

statistically significant

Step 6. State a clear conclusion that includes the p-value and whether you reject the null or fail to reject the null and why

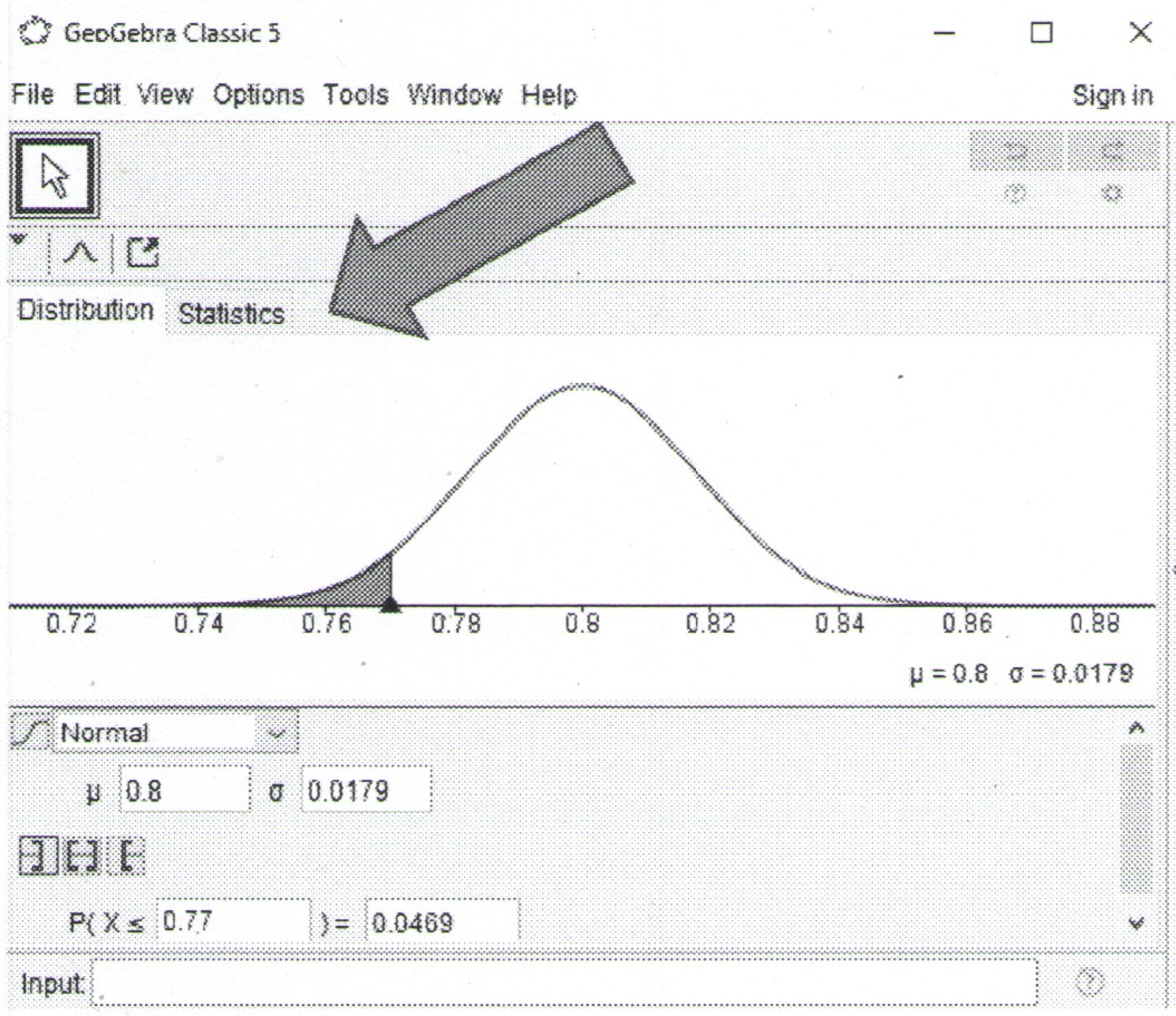
The p-value of .0469 is less than .05, so the result is statistically significant. We reject the null hypothesis because we have evidence to suggest that too few of its cars meet emission standards.

Test Statistic Addendum

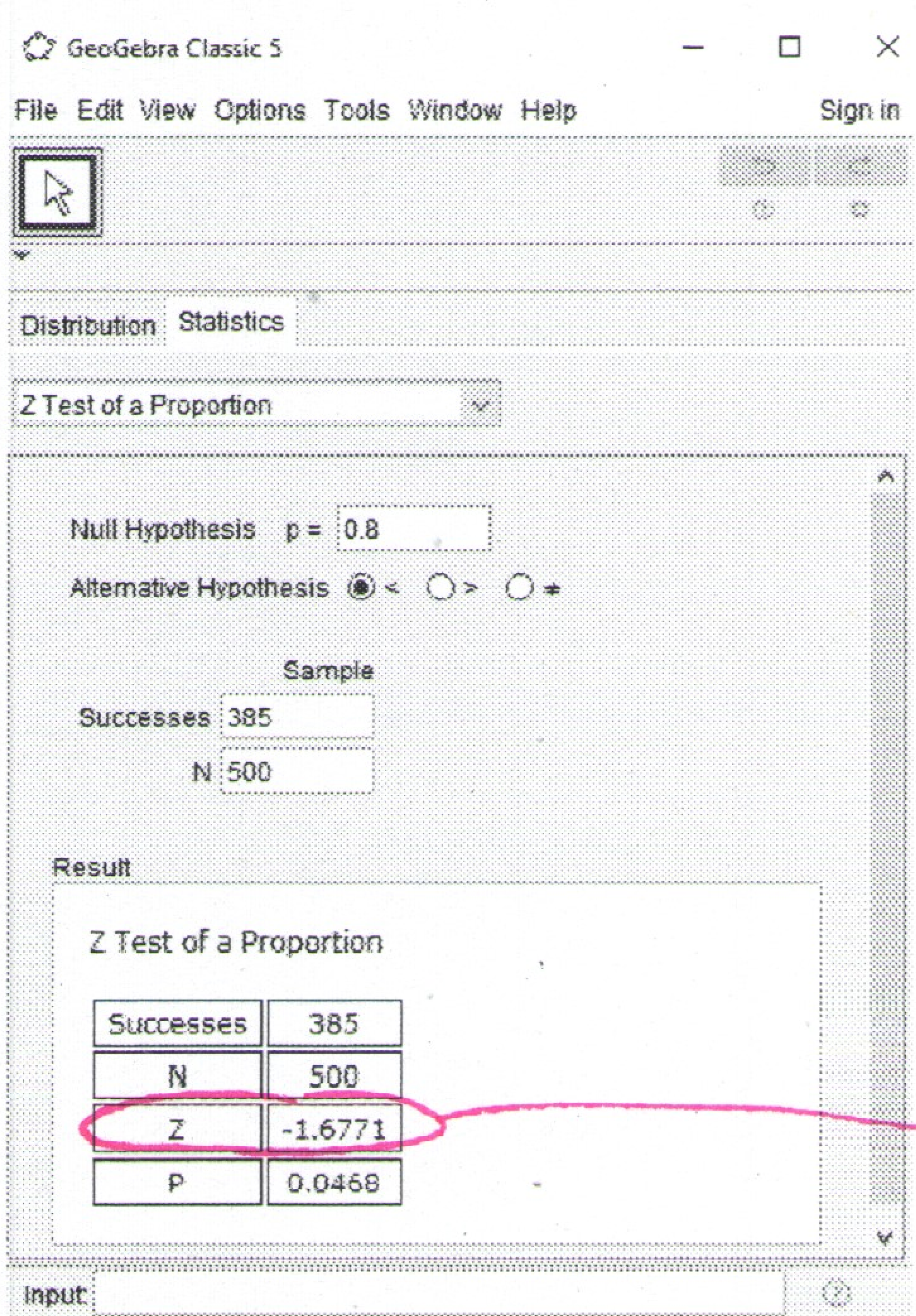
Use the test statistic feature of GeoGebra to find the test statistic and the p-value:

Example 3 Continued. Let's continue the emissions example. A company is concerned that too few of its cars meet pollution standards. They want to test whether less than 80% of their fleet meets emissions standards. They measure a random sample of 500 vehicles and 385 meet the standards. Is this evidence that too few of its cars meet the standards?

From the Normal Distribution, Click on the **Statistics Tab** (Test Statistics)



Use **Z Test of a Proportion** and enter the information.



test statistic

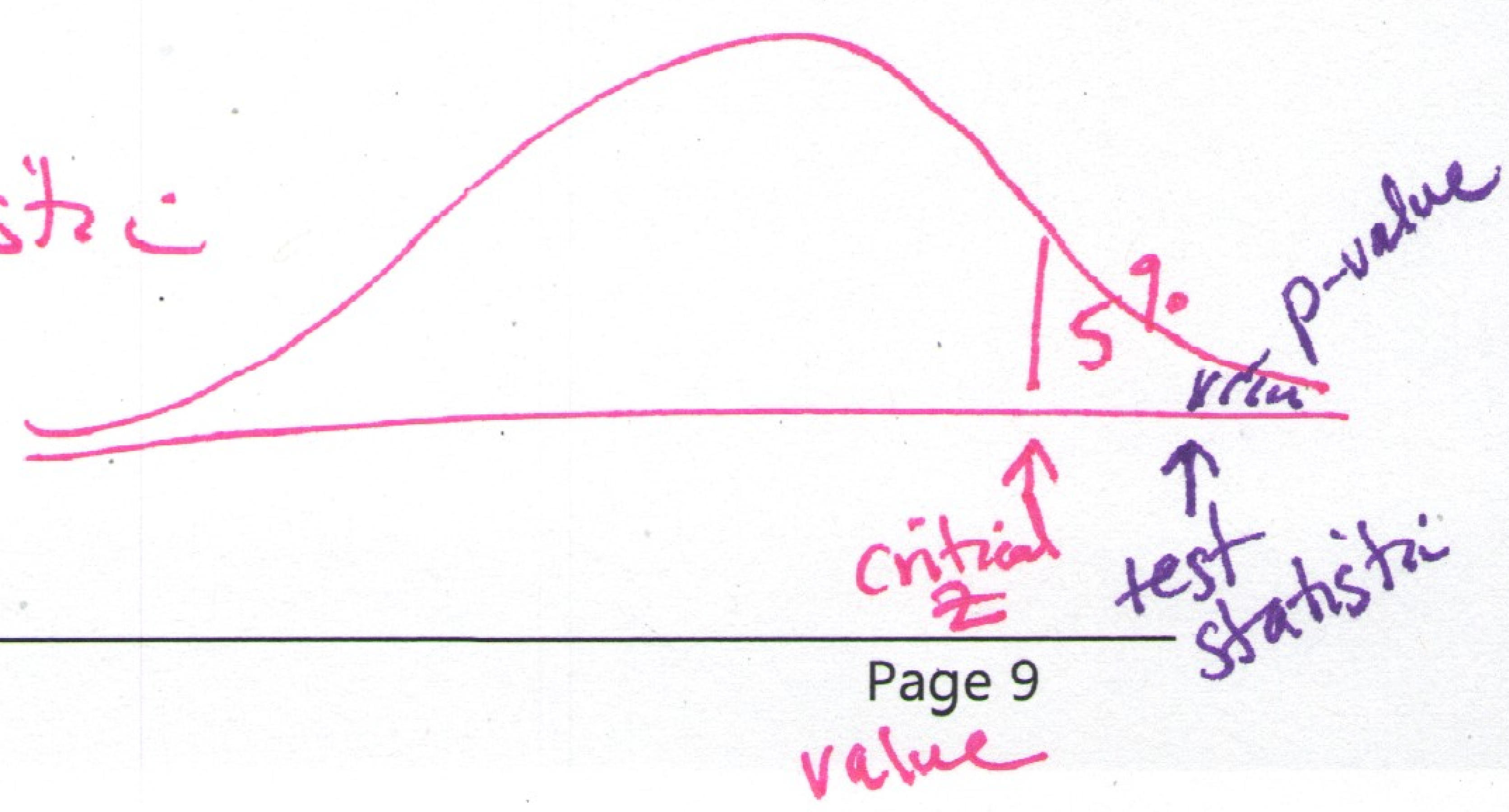
The test statistic is the Z-score of the p-value. Some instructors and books calculate the test statistic first and then the p-value. In Math 244 you will learn additional types of hypothesis tests.

We can either compare:

the p-value with the significance level, α

or

the test-statistic (Z) with a critical value Z_α



Example 4. Let's continue the poverty rate example. In 2014, the official poverty rate in the U.S. was 14.8%. A local official wants to test whether their county has a different poverty rate than the rest of the U.S. In a random sample of 2000 county residents, 13.3% were below the poverty level. Is this enough evidence to show that the county's rate is significantly different than the national rate?

$H_0: p = .148$

$H_A: p \neq .148$

(2-tailed test)

conditions
1. independence
2. random — need a representative random sample
3. 2000 < 10% of county
4. $2000(.148) = 296$ ✓
   $2000(.852) = 1704$ ✓

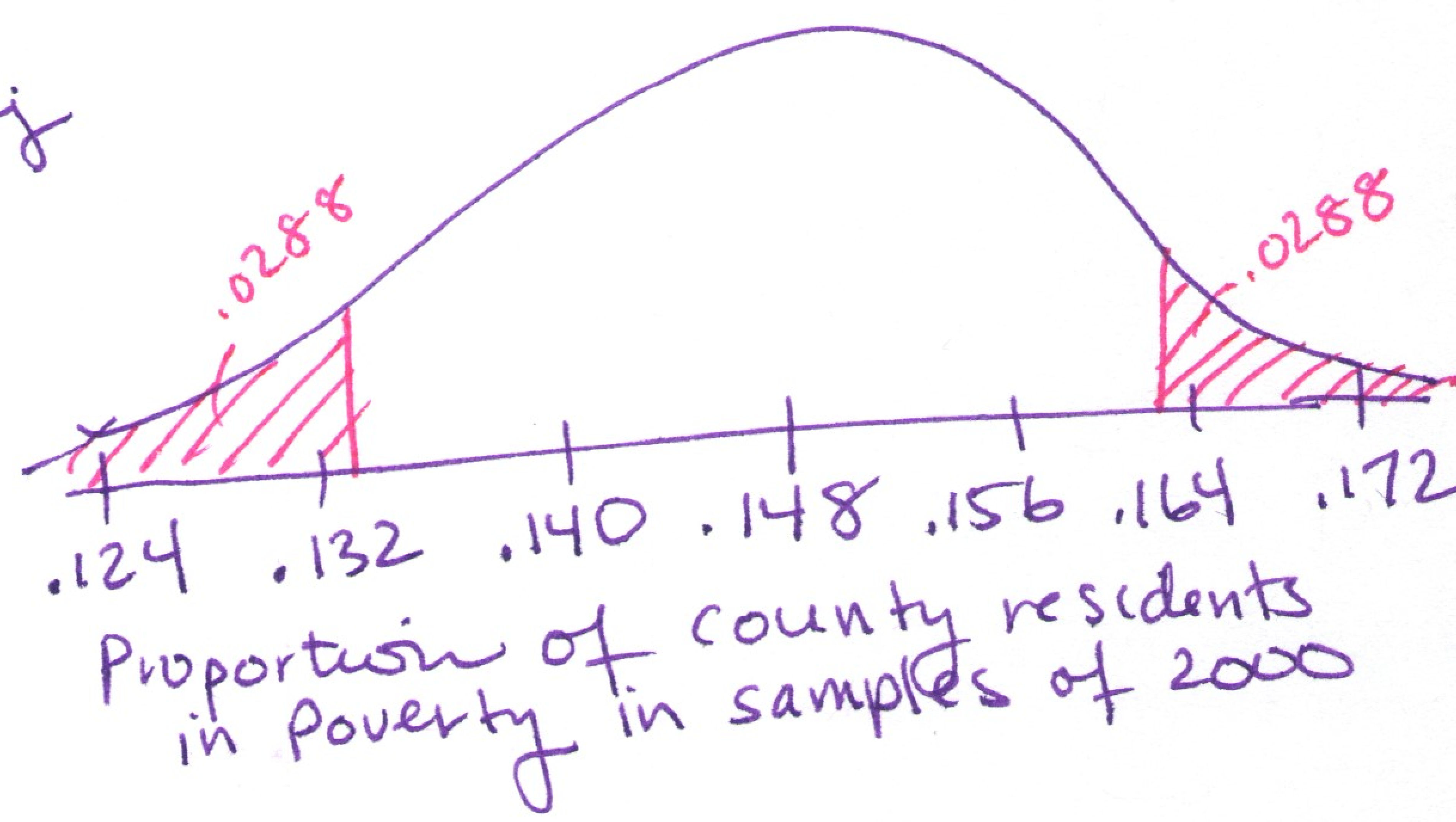
If Ho is true

$\hat{p} \sim N\left(.148, \sqrt{\frac{.148(.852)}{2000}}\right)$

$\approx .0079$

$\approx .008$ for drawing

$\hat{p} = .133$



p-value $= 2 \cdot P(\hat{p} \leq .133)$

$= 2 \cdot .0288$

$= .0576$

$.0576 \not< .05$

The p-value of .0576 is not less than .05 so the result is not statistically significant. We fail to reject the null and we do not have enough evidence to suggest that the poverty rate is different.

Find the 95% confidence interval for the proportion of county residents who are below the poverty level. How does this relate to the hypothesis test?

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$1 - .133 = .867$

$.133 \pm 1.96 \sqrt{\frac{.133(.867)}{2000}}$

$.133 \pm .0149$

$(.1181, .1479)$

$(12\%, 15\%)$

If the value for the null hypothesis is in the confidence interval then that supports not rejecting the null. In this case .148 is right on the edge, so it is very close. The p-value was just over .05 so they are similar results.

Practice 1. A company develops what it hopes will be better instructions for its customers to set up their smartphones. The goal is to have 96% of customers succeed. The company tests the new system on 400 people, of whom 376 were successful. Is this evidence that the new system fails to meet the company's goal?

$H_0: p = .96$ The company meets its goal of 96% succeeding

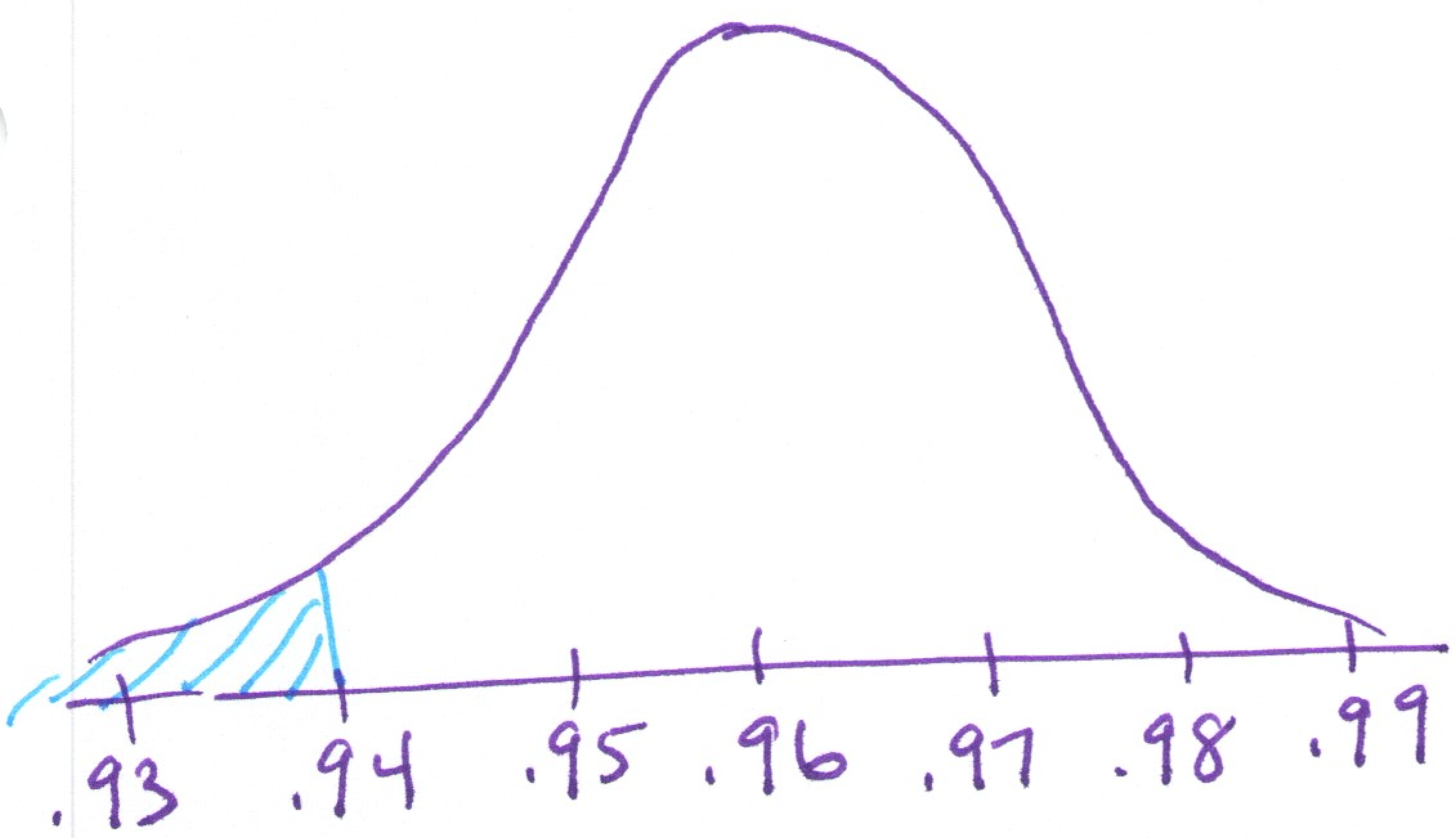
$H_A: p < .96$ The company fails to meet its goal

Conditions:

① independence
② random } if they use a random sample
③ <10% 400 is less than 10% of customers
④ $np = 400(.96) = 384 \geq 10$ ✓ $nq = 400(.04) = 16 \geq 10$ ✓

If Ho is true,

$p = .96$, $\sigma = \sqrt{\frac{.96(.04)}{400}} \approx .0098 \approx .01$ $N(.96, .0098)$



Proportion of customers who succeed in samples of 400.

$\hat{p} = \frac{376}{400} = .94$

p-value $= P(\hat{p} < .94)$
$= .0206 < .05$ significance level

The p-value is .0206, which is less than .05. There is evidence to suggest that the new system fails to meet their goal. We reject the null hypothesis.

Compare to 95% confidence interval:

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$.94 \pm 1.96 \sqrt{\frac{(.94)(.06)}{400}}$

$.94 \pm .0233$

$(.9167, .9633)$

$(92\%, 96\%)$

In this case the confidence interval does not agree with the hypothesis test because 96% is within the confidence interval. (Although it is on the edge!)

Practice 2. During the 2013 NFL season, the home team won 153 of 245 regular-season games. Is this enough evidence to conclude that there is a home-field advantage? (Hint: If there was not a home-field advantage, you would expect home teams to win 50% of the time, so the null is p=0.50)

$H_0: p = 0.50$ The home-team wins 50% of the time (There is no home-field adv.)

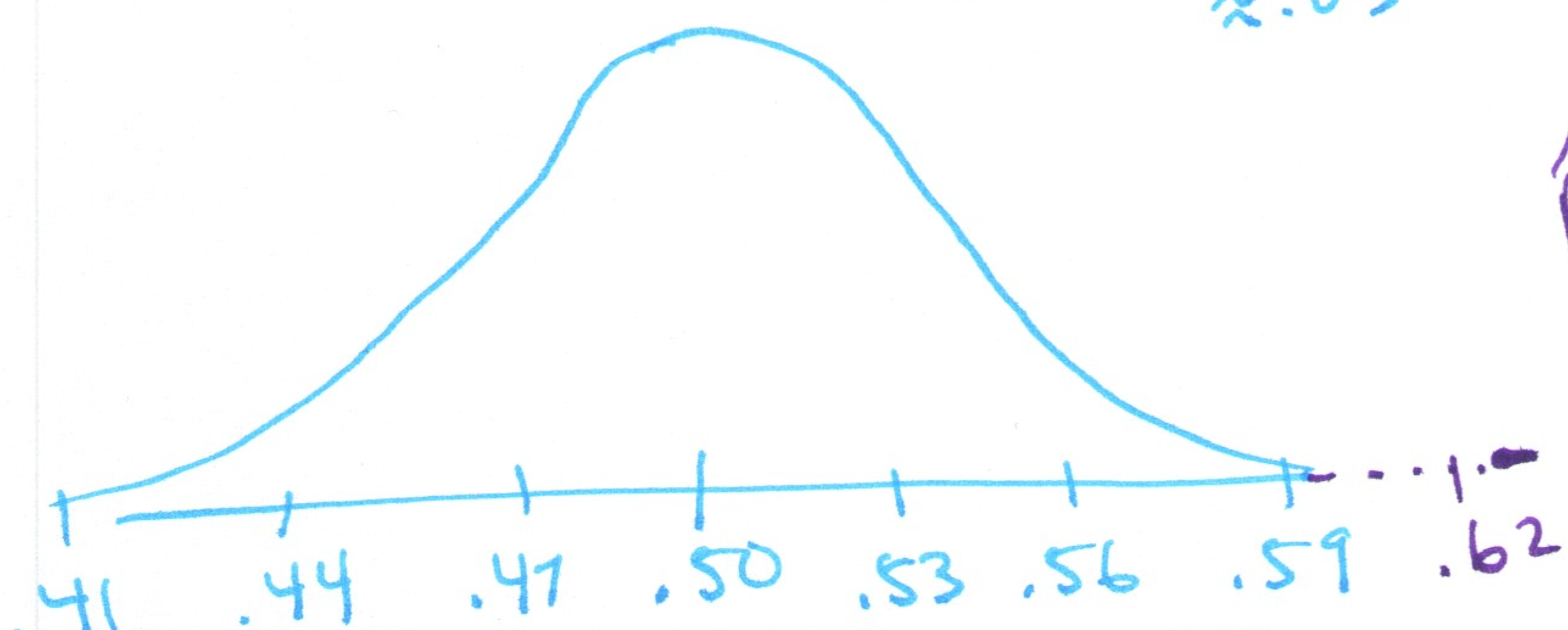
H_A: $p > .50$ The home team wins more than 50% of the time (There is a home-field advantage)

Conditions
① independence
② randomness
→ these are all the NFL games for one season so may be seasons are independent from each other and this is one random season.

③ <10%
One season is less than 10% of all

④ $np = 245(.5) = 122.5 \geq 10$ ✓ nq = same ✓

if Ho is true,
$p = .5$ $\sigma = \sqrt{\frac{(.5)(.5)}{245}} \approx .0319$
$\approx .03$



$\hat{p} = \frac{153}{245} = .6245$

p-value $= P(\hat{p} > .6245)$
≈ 0 $< .05$

The p-value is essentially zero, so we have enough evidence to suggest that there is a home-field advantage. We reject the null hypothesis.

Compare with 95% Confidence Interval

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$.6245 \pm 1.96 \sqrt{\frac{.6245(.3755)}{245}}$

$.6245 \pm .0606$

$(.5639, .6851)$
$(56\%, 69\%)$

The confidence interval says that the number of home-games won is from 56% to 69% with 95% confidence. 50% is not in the confidence interval so this agrees with our hypothesis test.

$\hat{p}$

Practice 3. A company is criticized because only 13 of 43 executive-level positions are held by women. The company counters that this is not unusual because only 40% of their employees overall are women (That's a separate issue). Is there enough evidence to suggest that the percentage of women in executive roles is less than 40% (the percentage of women overall)?

$H_0: p = .40$ The proportion of women in executive roles is the same as the company overall.

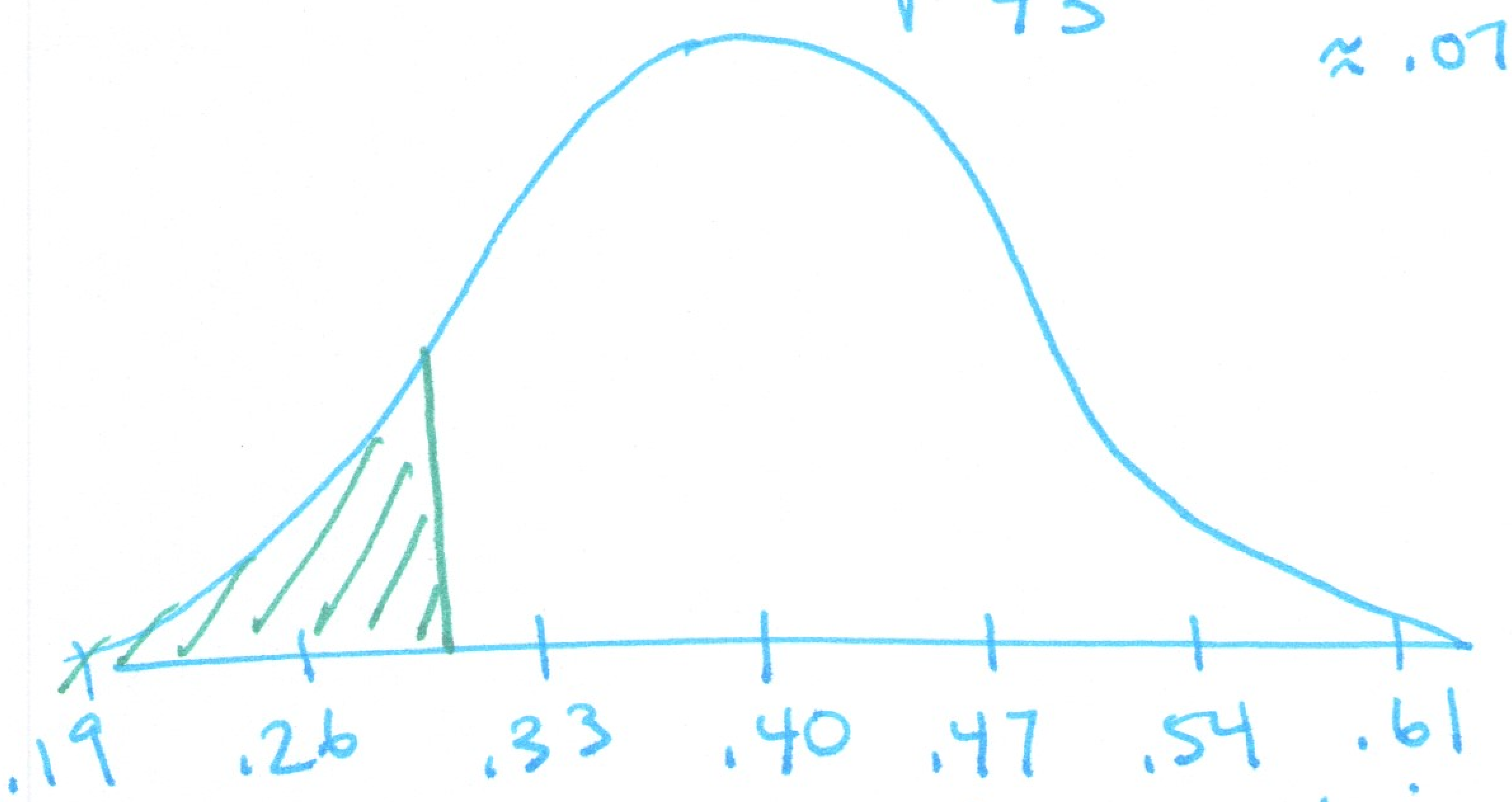
$H_A: p < .40$ There are fewer women in executive roles.

Conditions

1. independence - it's hard to say whether people in executive roles are independent - relationships are important
2. random - probably not
3. < 10% if 43 is less than 10% of the company
4. $np = 43(.40) = 17.2 \geq 10$ ✓ $nq = 43(.60) = 25.8 \geq 10$ ✓

if H_0 is true,

$p = .40$, $\sigma = \sqrt{\frac{.4(.6)}{43}} \approx .0747$ $\approx .07$ $N(.40, .0747)$



$\hat{p} = \frac{13}{43} = .3023$

p-value $= P(\hat{p} < .3023)$

$= .0955 \not< .05$

The p-value is .0955 which is not less than .05 so we fail to reject the null hypothesis. There is not enough evidence to suggest that the proportion of women in executive roles is less than in the company overall.

Compare with 95% Confidence Interval:

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$.3023 \pm 1.96 \sqrt{\frac{.3023(.6977)}{43}}$

$.3023 \pm .1373$

$(.165, .4396)$

$(17\%, 44\%)$

The confidence interval agrees with the hypothesis test because 40% is within the confidence interval.