Name: ______________________________

AP Statistics CED 6.1 Daily Video (Skill 1.A)

Introducing Statistics – Why Be Normal?

What Will We Learn?
How do we identify evidence for a claim?
How do we determine if the evidence for a claim is convincing?

Does Green = More Natural?
Many companies use the color green in the packaging of a product to suggest that it is healthier and more natural than other similar products.

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade.

Of the 30 students, _____ stated that the lemonade in the green cup tasted more natural.

Does Green = More Natural?
Do the data provide _________________ _______________that students associate the color green with being more natural?

What is the _____________________that students associate the color green with being more natural?

If the color of the cup had no effect on the subject's choice, _______________ of the students should have picked the green cup. However, in the study, ________________________________.

Does Green = More Natural?
What are the two explanations for the evidence that the student researchers found?

1. The color of the cups had _______ effect on the subject's choice and the researchers got a value of ________________ by _______________ alone.

2. Students __________________ the color green with being more natural.

To be convinced that Explanation #2 is correct, we need to know ___________________________ it would be to get a sample proportion of _______________ by chance alone. If it is unlikely to get _______________ by chance alone, we can rule out Explanation #1 and conclude that Explanation #2 is correct.

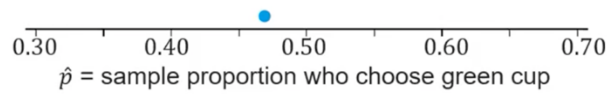
Does Green = More Natural?
How likely is it to get a sample proportion: $P(\hat{p} \geq 0.60)$ when $p = 0.50$ and $n = 30$? To answer this question, we can perform a ____________________.

Because we are assuming that true proportion of people who would choose the green cups 50/50 or _________, we can flip a coin _________________ to represent students choosing a cup at random

and let the __________________ of heads represent the _________________ of students who choose the green cup.
In the first set of 30 flips, the proportion of heads is $\hat{p}$ = 14/30 = __________________.
In a second set of 30 flips, the proportion of heads is $\hat{p}$ = 17/30 =_________________.
In a third set of 30 flips, the proportion of heads if $\hat{p}$ = 12/30 = ___________________.

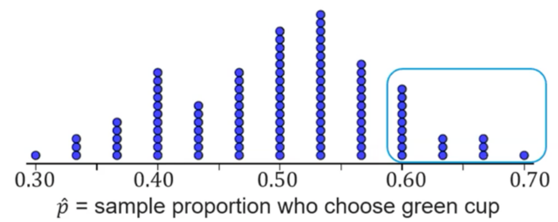
(Add the dots for the 2nd an 3rd simulation to the graph at the right.)

In the three simulations above, was there a sample that yielded a $\hat{p} \geq .60$ as was seen in the student's research? _____________ But, just three simulations is not enough to make a determinations, so we will run 100 trials of this simulation.



Does Green = More Natural?

Here are the results of 100 trials of this simulation. Based on the simulation, would it be surprising to get a sample proportion of $\hat{p}$ = 0.60 or greater by chance alone?
ANSWER: ____.



In 16 of the 100 trials, $\hat{p}$ was 0.60 ___________________. So, $P(\hat{p} \geq 0.60) \approx \frac{16}{100}$

Does Green = More Natural?

Let's revisit the two explanations for the evidence that the student researchers found in their study.
1. The color of the cups had no effect and the researcher got a value of $\hat{p}$ = 0.60 by chance alone.
2. Students associate the color green with being more natural.

Because it is ___________________ likely (_______________________________) to get a sample proportion of $\hat{p}$ = 0.60 or ______________ by chance alone when p = .50, we cannot rule out Explanation #1. *These data* ______________________ *provide convincing evidence that students at this school* ____ *with being more natural.*

What Should You Take Away?

How do we identify evidence for a claim?

- **Show that the _____________ of a study are consistent with the _______________.**

How do we determine if the evidence for claim is convincing?

- **Consider the ________ explanations for the ________________ (random chance, real effect).**
- **Estimate the _________________ of getting evidence as strong or _______________ than the observed evidence by _________________ alone.**
- **If you can eliminate the __________________ chance as a ______________________ explanation for the evidence, the evidence for the claim is ________________________.**

Name: ______________________________

AP Statistics CED 6.2 Daily Video 1 (Skill 4.C)

Constructing a Confidence Interval for a Population Proportion

What Will We Learn?

How do we identify an appropriate confidence interval procedure for a population proportion?
How do we verify the conditions for calculating a confidence interval for a population proportion?

Some Preliminary Information

There are two major types of inference that you will learn about in Units 6 – 9:

- Confidence Intervals
  - These are used to ______________ the value of a population _____________, such as a population ______________, mean or slope.
  - We use an ______________ of values – rather than a single value – to ______________ a parameter to account for sampling __________________.
- Significance Tests
  - These are used to ___________________ about the value of a population parameter, such as population proportion, _________ or slope.
  - We use _________________________ to decide whether the evidence supporting a claim is likely or unlikely to happen by ______________________.

In Unit 6, we will focus on ______________________ data that can be summarized by calculating the proportion of _________________ in a sample or treatment group.

Topics 6.2 – 6.3 focus on __________________________ for a population proportion.
Topics 6.4 – 6.7 focus on __________________________ for a population proportion.
Topics 6.8 – 6.11 focus on confidence intervals and significance tests for a ________________ in population proportions.

Example: Verifying Signatures

In some cities and states, citizens can process new laws to be voted on in an election. To get a proposition on the ballot, backers of the proposition must gather a specified number of signatures from registered voters in the city.

Supporters of "Proposition 100" submitted 9388 signatures in support of this proposition. City administrators randomly selected 500 of these signatures to verify they came from registered voters in the city. Of the 500 signatures administrators were able to verify that 364 of the signatures were from registered voters in the city. **Calculate and interpret a 95% confidence interval** for the proportion of all submitted signatures that are from registered voters in the city.

Identifying the Procedure

There are many difference confidence interval procedures that you will learn this year. Some involve one sample and others involve two samples. Some involve estimating proportions, some involve estimating ____________ and some involve estimating ___________.

When the goal is to estimate the _____________________ of _______________ in a single population, we use a ____.

Checking the Conditions

After identifying the correct procedure, you must verify that the ________________ for using that procedure are met.

In general, you should always be checking for:

- ______________________ in the methods used to collect the data, and
- that the appropriate ___________________ distribution has the correct shape.

Checking the Conditions for a one-sample *z* interval for a __________________ proportion.

To check for ______________________:

1. The data are collected using a ________________ sample from the population.
2. When sampling _______________ replacement, the ______________ ________ is less than or equal to __________ of the population size.

To check that the shape of the ____________________ distribution is approximately normal:

3. Both _______________ and ________________. That is, the number of ______________ in the sample and the number of _____________ in the sample are both at least 10.

*Note that we have $\hat{p}$ there instead of p, that is because in the _________________ we do not know the truth about the population, that is what we are trying to figure out, so we use $\hat{p}$.

Checking the Conditions

Supporters of "Proposition 100" submitted 9388 signatures (that's our _________________ size) in support of this proposition. City administrators randomly selected 500 of these signatures (that's our _________________ size) to verify they came from registered voters in the city. Of the 500 signatures administrators were able to verify that 364 of the signatures were from registered voters in the city (so our value of $\hat{p}$ is ________/________). Check if the conditions for calculating a confidence interval are met.

1. The signatures were ________________ selected. ☐
2. The sample size (______) is less than _______ of the ________________ size (9388). ☐
3. $n\hat{p}$ = __________________ and $n(1 - \hat{p})$ = _________________. ☐

The conditions have been _________.

*Note: 364 is the number of ______________ and 136 is the number of ______________.

Check the Conditions, Part 2

Suppose the city administrators were in a hurry and only verified the first 40 signatures submitted. Of these 40, 38 were from registered voters in the city. Using these data, would it be appropriate to calculate a confidence interval for the proportion of all submitted signatures that are from registered voters in the city?

1. The signatures were _______________________ selected. ☐
2. The sample size of _______ is less than ________ of the population size. ☐
3. $n\hat{p}$ = ______________________ but $n(1 - \hat{p})$ = __________________. ☐

The conditions have ______ been met, so it is _______ appropriate to proceed.

What Should We Take Away?

How do we identify an appropriate confidence interval procedure for a population proportion?

When estimating the ____________ of successes in a population, use a ___________________ ___________________________

How do we verify the conditions for calculating a confidence interval for a population proportion?

- **The data is collected using a _____________ sample from the population.**
- **When sampling __________ replacement, the sample size is $\leq$ 10% of the population.**
- **Both _______________ and _________________.**

Name:______________________________

AP Statistics CED 6.2 Daily Video 2 (Skill 3.D)
Constructing a Confidence Interval for a Population Proportions

What Will We Learn?

How do we determine the margin of error when estimating a population proportion?
How do we calculate a confidence interval for a population proportion?

Example: Verifying Signatures

Supporters of "Proposition 100" submitted 9388 signatures in support of this proposition. City administrators randomly selected 500 of these signatures to verify they came from registered voters in the city. Of the 500 signatures administrators were able to verify that 364 of the signatures were from registered voters in the city. Calculate and interpret a 95% confidence interval for the proportion of all submitted signatures that are from registered voters in the city.

***NOTE:** In the previous video, we identified the procedure and checked the conditions.

Some Preliminary Information…

Based on what we learned in Unit 5, the sample proportion _______ is an unbiased estimator of the population proportion ________. However, because of sampling variability, the value of ________ will almost never ___________ the value of p. In other word, we should have ________ confidence that the value of _______ correctly ______________ the population proportion. To increase our confidence that our _____________ is correct, we use an interval of values (a ________________ ________________) as our estimate rather than a single value (a ____________ ____________.)

Calculating a Margin of Error

In AP Statistics, confidence intervals have the form:

CI = _______________________ ± __________________________

The margin of error described how much the value of a ____________ ____________ is likely to vary from the value of the corresponding _____________________ _____________.

The margin of error is determined by two factors:

- How much the ______________ typically varies from the _________________.
- How confident we want to be in our estimate

margin of error = (___________________) (______________________________)

Calculating the Margin of Error

margin of error = (critical value) (_________________________)

The standard error of a statistic is an estimate of the ___________________________ of the _______________________ distribution of the statistic. Typically, to find the true standard deviation, you need to know information about the ____________________. In this case, we will only know information about the __________________.

From Topic 5.5, the standard deviation of the sampling distribution of $\hat{p}$ is: $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Because we don't know the value of ______, we replace p with $\hat{p}$ to get the ________________________________. $SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Calculating the Margin of Error

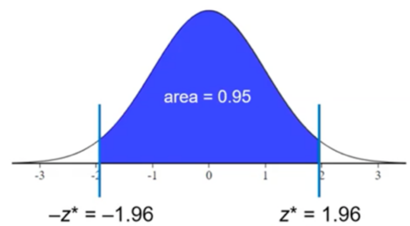
margin of error = (____________________)(standard error of statistic)

The ___________ __________ is a multiplier that makes the margin of error large enough to give a specific amount of ___________________ that the interval _______________ the value of the _________________. For confidence intervals for a ____________________, the critical values represent the boundaries encompassing the _______________ *C*% of the ________________ normal distribution, where *C*% is the ______________________ confidence level for a proportion.

Calculating the Margin of Error

Margin of error - (____________ __________)(standard error of statistic)

In the verifying signature example, we are asked to construct a 95% confidence interval. To find the ____________ value *z** for a 95% confidence interval, find the boundaries encompassing the ____________ 95% of the _____________ normal distribution.



The critical value of z* = __________.

Calculating the Confidence Interval

CI = ______________ ± ___________________________; in other words,

CI = ______________ ± (____________________) (________________________________)

For our example about verifying signatures, $\hat{p}$ = _______/_______ =_______

Therefore:

$$\text{CI} = \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

=________________________________

=________________________________

=________________________________

***NOTE: Stop** and find the information you need to calculate the CI on the formula sheet!!

Calculating the Confidence Interval

The information you need to calculate the confidence interval is available on the formula sheet. Stop find this on your formula sheet.

Calculating the Number of Valid Signatures

We can use confidence intervals for a population proportion to estimate the ____________ of successes in a population. Because there were 9388 signatures in the population, a 95% confidence interval for the ______________ of valid signatures is:

9388(0.689) to 9388(0.767) ≈ __________ to __________ valid signatures.

What Should We Take Away?

How do we determine the margin of error when estimating a population proportion?

margin of error = (__________________) (___________________________)

How do we calculate a confidence interval for a population proportion?

CI = __________________ ± ___________________

AP Statistics CED 6.2 Daily Video 3 (Skill 3.D)

Construction a Confidence Interval for a Population Proportion

What Will We Learn?

How do we determine the minimum sample size that will achieve a given margin of error?

What Proportion of Students Bring Lunch?

The director of food services in a large high school district wants to estimate the proportion of students who always bring lunch from home. What is the smallest sample size that will result in a margin of error of at most 0.05 with 95% confidence?

*Note: We have NOT taken our sample yet!

For a one-sample *z* interval for a population proportion:

$$\text{margin of error} = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

What Proportion of Students Bring Lunch?

We want to solve the following inequality for *n*:

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le 0.05$$

Because we want a 95% confidence interval, z* = ___________.

What value do we use for $\hat{p}$?

- Use a guessed value for $\hat{p}$ based on a pilot study or _____________ _______________.
- Use $\hat{p}$ = _____ in order to find an ___________ bound for the sample size.

What Proportion of Students Bring Lunch?

Using $\hat{p}$ = 0.5 we can now solve for *n*: $1.96\sqrt{\frac{0.5(1-0.5)}{n}} \le 0.05$ (Solve along with the video)

*NOTE: When determining the smallest sample size, you always round _________!!

What Proportion Read the Comics?

The editor of a newspaper wants to estimate the proportion of subscribers who read anything from the comics section. What is the smallest sample size that result in a marge of error of at most 0.03 with 90% confidence?

To do this you will need: $\hat{p}$ = ______ and;

*z** for 90% = _____________

$$\_\_\_\_\sqrt{\frac{0.5\,(1-0.5)}{n}} \le \_\_\_\_$$

(Fill in the blanks to the right and solve)

_________ ≤ *n* ➡ The editor needs to survey at least __________ subscribers.

What Should We Take Away?

How do we determine the minimum sample size that will achieve a given margin of error?

margin of error = ____________________

Use a guessed value of $\hat{p}$ or us $\hat{p}$ = __________ to find an upper bound for the required sample size.

***NOTE:** This procedure only works when we are trying to estimate a _____________ ______________.

Name___________________________

AP Statistics CED 6.3 Daily Video 1 (Skill 4.B)

Justifying a Claim Based on a Confidence Interval for a Population Proportion

What Will We Learn?

How do we interpret a confidence interval for a population proportion?

How do we justify a claim based on a confidence interval for a population proportion?

Verifying Signatures

Supporters of "Proposition 100" submitted 9388 signatures in support of this proposition. City administrators randomly selected 500 of these signatures to verify they came from registered voters in the city. Of the 500 signatures administrators were able to verify that 364 of the signatures were from registered voters in the city. Calculate and interpret a 95% confidence interval for the proportion of all submitted signatures that are from registered voters in the city.

***Note:** In previous videos we have already 1) identified the procedure, 2) checked the conditions and 3) done the calculations for this interval.

Interpreting the Confidence Interval

In general, here is how to interpret a confidence interval for a ________________ __________________:

"We are *C*% confident that the interval from ______ to ______ captures the [population parameter]. (***Note:** We always interpret things ____ ___________ in AP Statistics!)

From Topic 6.2, Video 2, the 95% confidence interval is ___________ to ___________.

We are ___________ confident that the interval from _______________________ captures the proportion of ____

____.

Justifying a Claim

According to city administrators, propositions need at least 6000 valid signatures to qualify to be on the ballot in the next election. Do the supports of Proposition 100 have enough signatures?

Because the supporters submitted ________ signatures, the required proportion of valid signatures is _________________________ = ___________.

Name____________________________

Because all of the values in the confidence interval (0.689 to 0.767) are greater that 0.639, there is ________________ ____________________ that the supporters of Proposition 100 have enough signatures.

Will Proposition 100 Pass?

Now that Proposition 100 is on the ballot, a local television news program wants to estimate the proportion of registered voters who plan to vote for the proposition. Based on a random sample of 500 registered voters in the city the new program reported a 95% confidence interval of 0.518 ± 0.044.

(a) Interpret the confidence interval. (Remember the sentence from above!)

(b) Based on the interval, is there convincing evidence that a majority of registered voters in the city plan to vote for Proposition 100?

What Should We Take Away?

How do we interpret a confidence interval for a population proportion? (copy sentence)

____

____

How do we justify a claim based on a confidence interval for a population proportion?

- **If ______ the values of the confidence interval are consistent with the claim, there is ________________________________ for the claim.**
- **If ________________ of the values in the confidence interval are inconsistent with the claim, there is __________ convincing evidence for the claim.**

Name________________________________

AP Statistics CED 6.3 Daily Video 2 (Skill 4.B)

Justifying a Claim Based on a Confidence Interval for a Population Proportion

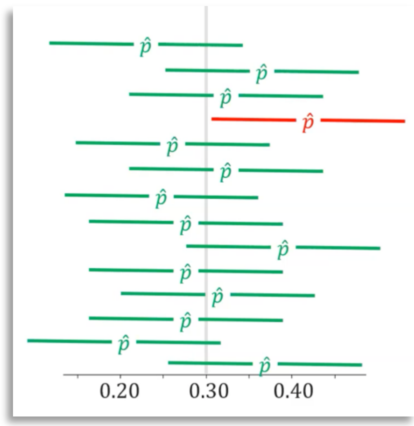
What Will We Learn?

How do we interpret the confidence level of a confidence interval for a population proportion?

How do the sample size and confidence level affect the margin of error for a confidence interval for population proportion?

Sampling From a High School

(Many Samples)



In a high school with 2000 student, 30% of the students have a driver's license. Suppose we select a random sample n = 50 students from the high school, as each student in the sample if he or she has a driver's license, and calculate a 95% confidence interval for the proportion of all student with a driver's license.

So far, __________ = ________ of the intervals have captured p = 0.30, the proportion of all students with a driver's license. If we continued this process, about 95% of the "95% __________________ ____________ would capture p = ___________. So if the confidence level is ______, you know that if you took lots of samples and used those samples to make lots if interval that about __________ of those intervals would end up capturing the ____________________________________.

Interpreting Confidence Level

In repeated random sampling with the ________________________________, approximately C% of "*C*%" confidence intervals will ______________________ the population proportion.

"If we take ________ random samples of ____________ from the population of students at __________ high school and use each sample to construction a 95% confidence interval for the ______________________ of all students with a driver's license, about _________ of those intervals would ______________ the population proportion."

Interpreting Confidence Level

The confidence level describes what happens in ________________________________. It does ______ give the probability that a ____________________ interval captures the population proportion. (You will see this on the AP exam as a multiple-choice question!!)

Factors That Affect the Margin of Error

Recall that confidence intervals in AP Statistics have the following structure:

CI = ______________________ $\pm$ _____________________________________

The width of a confidence interval is _________________ the margin of error.

We generally prefer narrow confidence intervals (more precision), so we want the ______________________________ to be small.

Name________________________________

Factors That Affect the Margin of Error

There are two common ways to decrease the margin of error. Assuming everything else remains the same, the margin of error will be _______________ when…

1. The sample size is __________________, because in the formula:

$$\text{margin of error} = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

For a _____________________________, the width of the interval is proportional to $\frac{1}{\sqrt{n}}$. This means that _______________ the sample size will cut the margin of error in _______.

2. The confidence level is smaller, because in the formula:

$$\text{margin of error} = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

The critical value for a 90% confidence (z^* = _________) is smaller than the critical value for 95% confidence (z^* = _________), making the 90% CI narrower than the 95% CI.

What Should We Take Away?

How do we interpret the confidence level of a confidence interval for a population proportion?

In repeated random sampling with the ______________________________, approximately *C*% of *C*% confidence intervals will ___________ the __________________ proportion.

How do the sample size and confidence level affect the margin of error for a confidence interval for population proportion?

Assuming everything stays the same:

- **_________________ the sample size will __________________ the margin of error.**
- **_________________ the confidence level will __________________ the margin of error.**

Name____

AP Statistics CED 6.3 Daily Video 3 (Skill 4.D)

Justifying a Claim Based on a Confidence Interval for a Population Proportion

What Will We Learn?

How do we construct and interpret a confidence interval for a population proportion?

2010 B #4

A husband and wife, Mike and Lori, share a digital music player that has a feature that randomly selects which song to play. A total of 2,384 songs were loaded onto the player, some by Mike and the rest by Lori. Suppose that when the player was in the random-selection mode, 13 of the first 50 songs selected were songs loaded by Lori.

(a) Construct and interpret a 90 percent confidence interval for the proportion of songs on the player that were loaded by Lori.

(b) Mike and Lori are unsure about whether the player samples the songs with replacement or without replacement when the player is in random-selection mode. Explain why this distinction is not important for the construction of the interval in part (a).

2010 B #4 part (a)

(a) Construct and interpret a 90 percent confidence interval for the proportion of songs on the player that were loaded by Lori.

1. Define the parameter you are trying to estimate:

   ____

   ____

2. Identify the procedure that will be used to create the confidence interval:

   ____

3. Check the condition to make sure the chosen procedure is valid:

- ____ ☐
- ____ ☐
- ____ ☐

The conditions ____________________ met.

2010 B #4 part (a)

(a) Construct and interpret a 90 percent confidence interval for the proportion of songs on the player that were loaded by Lori.

4. Find $\hat{p}$ = _____________
5. Calculate the CI: =________ $\pm$ _________________________

   =________ $\pm$ __________

   =________ to ________

Interpret the interval: ____

____

STATS MEDIC

Name____

2010 B #4 part (b)

(b) Mike and Lori are unsure about whether the player samples the songs with replacement or without replacement when the player is in random-selection mode. Explain why this distinction is not important for the construction of the interval in part (a).

When sampling without replacement, the formula for the standard deviation of the sampling distribution of $\hat{p}$ will ________________________ the actual value of the standard deviation, unless the _____________ size is small relative to the _____________________ size.

(See topic 5.5) In this case, the sample size of ________ is less than ________ of the population size of 2384, so we can use the formula as if we were sampling with replacement.

What Should We Take Away?

How do we construct and interpret a confidence interval for a population proportion?

Make sure to:

- Define the ________________ you are trying to estimate.
- Identify the _______________________ you are using.
- Verify that the ____________________ for the ___________________ are met (with evidence!).
- Calculate the _______________________________.
- Interpret the interval ____________________________. (You do not need to interpret the confidence *level* unless specifically asked.)

Name______________________________

AP Statistics CED 6.4 Daily Video 1 (Skill 1.F)

Setting Up a Test for Population Proportion

What Will We Learn?

How do we state a null hypothesis in a test for a population proportion?

How do we state an alternative hypothesis in a test for a population proportion?

Does Green = More Natural?

Many companies use the color green in the packaging of a product to suggest that it is healthier and more natural than other similar products.

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural.

Null Hypothesis:

In a statistical test, the ________ hypothesis is often a claim of "_____ difference" or "_____ change."

In the lemonade example, the null hypothesis is that there is ______ _______________ in the _______________ of students who would choose the green cup and the proportion who would choose the white cup. In other words, _______ of the students would choose the green cup. In symbols:

_______: p = 0.50 where p is the proportion of ______ students at ______ ________ who would choose the ___________cup. Until we have ____________________ evidence otherwise, we ______________ H_0 is ______________.

Alternative Hypothesis

In a statistical test, the _________________ hypothesis is the claim that we ________ to support with evidence from the ___________ collected.

In the lemonade example, the researchers wanted to know if the students associated the color ________ with being more ____________. So, the alternative hypothesis is that more than _______ of the students would choose the __________ cup when selecting the lemonade that tasted more ____________. In symbols:

H_0: ____________

H_a: ____________, where p = the ________________ of ________ students at this school who would choose the ____________ cup.

STATS MEDIC

Name______________________________

Stating Hypotheses

For hypotheses about a parameter:

The __________ always contains a strict ___________ reference.
The ________________ always contains a strict ________________ reference.

- When the inequality is _____ or ______, the alternative is called _________________.
- When the inequality is ________, the alternative is called ____________________.
- The choice of the _________________ is determined by the _________________ of interest and should be stated _______________ data collection begins.

* Never include a _______________ (such as $\hat{p}$) in the hypotheses!! Hypotheses are always about the ___________________ and therefore should always be a ____________________.

* Remember to _____________ the parameter in terms of the _____________________.

Stating Hypotheses

A newspaper reports that 40% of the adults in the U.S. would say football is their favorite sport. The mayor of a town wonders if the proportions of adult in the town who would say football is their favorite sport differs from the national proportion. To Investigate, the mayor will select a random sample of 100 adults in the town and ask each adult to name his or her favorite sport.

Which inequality sign?
differs: __________
more: __________
less: __________

State appropriate hypotheses for the mayor's study.

___________ = ______________
___________ = ______________

Verb tense MATTERS!!
Present Tense refers to ________________
Past Tense refers to _____________

Where p =
____

What Should We Take Away?

How do we state the null hypothesis in a test for a population proportion?

H_0: ____________ [hypothesized proportion]

where p = ____________________________ defined in ________________________

How do we state the alternative hypothesis in a test for a population proportion

H_a: _________ [hypothesized proportion]

OR

H_a: _________ [hypothesized proportion]

OR

H_a: _________ [hypothesized proportion]

Name_______________________________

AP Statistics CED 6.4 Daily Video 2 (Skill 4.C)

Setting up a Test for a Population Proportion

What Will We Learn?

How do we identify an appropriate significance test procedure for a population proportion?
How do we verify the conditions for performing a significance test for a population proportion?

Does Green = More Natural?

Many companies use the color green in the packaging of a product to suggest that it is healthier and more natural than other similar products.

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural.

In the previous video, we stated the hypotheses:

H_0: $p = 0.50$

H_a: $p > 0.05$, where p = the proportion of **all students at this school** who would choose the green cup.

Identifying the Procedure

When the goal is to test claim about the ________________ of successes in a _____________ population, we use a ____.

Checking the Conditions

After identifying the correct _________________, you must verify that the _________________ for using that procedure are met. In general, you should always be checking for:

- ___________________ in the methods used to _____________ the data, and
- that the appropriate _______________ distribution has the correct _________________.

* NOTE: Details of conditions vary from procedure to procedure!!

Checking the Conditions

Here are the conditions for a ___________________ for a ____________________ proportion.

To check for independence:

1. The data are collected using a _____________ sample from the population.
2. When sampling ____________ replacement, the sample size is less than or equal to _____________ of the population size.

To check that the shape of the sampling distribution is __________________ normal:

3. Both $np_0 \geq 10$ and $n(1 - p_0) \geq 10$, where p_0 is the proportion specified by the ___________

*Note that condition #3 differs for _______________________ and _______________________!

Name______________________________

Checking the Conditions

In the lemonade study, two student researchers randomly selected 30 students from their school. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural. Check if the conditions for performing a significance test are met.

1. The students were _____________ selected ☐

2. We assume that the ____________ size (30) is less than or equal to ________ of the total number of students at the school. ☐

3. np_0 = __________ = _____ ≥ _____ and $n(1 - p_0)$ = _____________ = _____ ≥ _____ ☐

*Note: We are not using the ______________ successes and failures from the sample, but we are using the ______________ number of successes and failures if the __________ hypothesis were true!

The _______________ are met.

Checking Conditions:

A newspaper reports that 40% of the adults in the U.S. would say football is their favorite sport. The mayor of a town wonders if the proportions of adult in the town who would say football is their favorite sport differs from the national proportion. To Investigate, the mayor will select a random sample of 100 adults in the town and ask each adult to name his or her favorite sport. Check if the conditions for performing a significance test are met.

1. ____. ☐

2. ____

____. ☐

3. ____. ☐

Therefore: ____

What Should We Take Away?

How do we identify an appropriate significance test procedure for a population proportion?

When testing a claim about the _______________ of successes in a population, use a ____.

How do we verify the conditions for performing a significance test for a population proportion?

1. **The data is collected using a ______________ ______________ from the population.**
2. **When sampling ______________ __________________, the sample size is __________ than or ____________ to ________ of the population size.**
3. **Both ________________ and ________________, where p_0 is the proportion specified by the ____________ ______________________.**

Name______________________________

AP Statistics CED 6.5 Daily Video 1 (Skill 3.E)

Interpreting *p*-Values

What Will We Learn?

How do we calculate an appropriate test statistic in a test for a population proportion?
How do we calculate a *p*-value in a test for a population proportion?

Does Green = More Natural?

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural.

From previous videos:

- H_0: p = 0.50 vs H_a: p > 0.05, where p = the proportion of all students at this school who would choose the green cup.
- Conditions are met.

Calculating a Test Statistic

In our lemonade study, $\hat{p} = \frac{18}{30} = 0.60 > 0.50$

This is evidence for H_a because ___________________________.

We want to know _______ ______________ it is to get evidence for H_a this ____________ or _______________ by chance alone ___________ H_0 is ____________.

After verifying the ______________ are met, calculate the _________________ test statistic.

$$\text{standardized test statistic} = \frac{\text{sample statistic} - \text{null value of the parameter}}{\text{standard deviation of the statistic}}$$

Calculating a Test Statistic

For a ____, the standardized test statistic is:

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

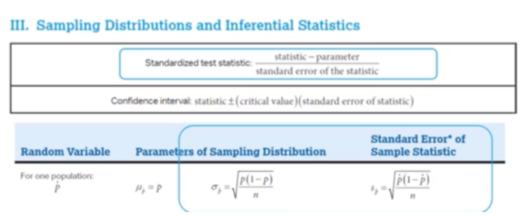
For our data it looks like this:

$$z = \frac{0.60 - 0.50}{\sqrt{\frac{0.50(1-0.50)}{30}}} = 1.10$$

Where p_0 is the value of p specified the null hypothesis.

Calculating a Test Statistic

The information that you need to calculate the standardized test statistic is available on the formula sheet. Locate that now!!



III. Sampling Distributions and Inferential Statistics

Standardized test statistic: $\frac{\text{statistic} - \text{parameter}}{\text{standard error of the statistic}}$

Confidence interval: statistic $\pm$ (critical value)(standard error of statistic)

Random Variable	Parameters of Sampling Distribution		Standard Error* of Sample Statistic
For one population: $\hat{p}$	$\mu_{\hat{p}} = p$	$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$	$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Name______________________________

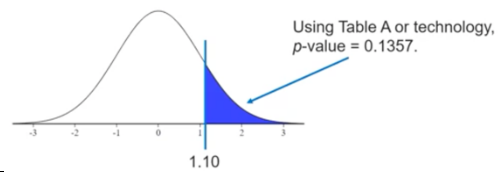
Calculating a *p*-value

Once we have calculated the __________________ test statistic, use the __________________ normal distribution to calculate the *p*-value. The *p*-value is the _________________ of _______ for the ________ distribution that are ___ __________ or _________ ___________ than the observed value of the ________ _____________ in the direction of the ____________________.

Calculating a *p*-value

Because our alternative hypothesis is _____________, we want to find $P(z \geq$ _______)

Start by drawing a picture and then use Table A or technology to determine *p*-value = 0.1357



Favorite Sport

A newspaper reports that 40% of the adults in the U.S. would say football is their favorite sport. The mayor of a town wonders if the proportions of adult in the town who would say football is their favorite sport differs from the national proportion. To Investigate, the mayor selected a random sample of 100 adults in the town and found that 29 named football as their favorite sport. Calculate the standardized test statistic and *p*-value.

From previous videos:

- H_0: p = ________ vs H_a: $p \neq$ ________, where p = the proportion of _________________ _______________ who would say football is their favorite sport.
- Conditions are ___________.

Calculating the Test Statistic

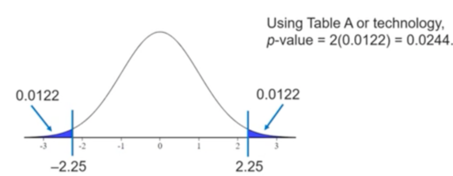
In this context, $\hat{p}$ = __________ = _________. Which means the test statistic would be:

$$z = \frac{\underline{\quad} - \underline{\quad}}{\sqrt{\frac{\underline{\quad}(1-\underline{\quad})}{\underline{\quad}}}} = \underline{\qquad\qquad}$$

Calculating the *p*-value

Because our alternative hypothesis is H_a: $p \neq$ _______ we want to find: $P(z \leq -2.25) + P(z \geq 2.25)$ = ___________.

(Because ≠ means we ADD both tails.)



What Should We Take Away?

How do we calculate an appropriate test statistic in a test for a population proportion?

$$z = \frac{\underline{\quad} - \underline{\quad}}{\sqrt{\frac{\underline{\quad}(1-\underline{\quad})}{\underline{\quad}}}}$$

How do we calculate a *p*-value in a test for a population proportion?

- **If H_a: __ > __, *p*-value = $P(z \geq$ ___________ test statistic)** (ONE TAILED)
- **If H_a: __ < __, *p*-value = $P(z \leq$ ___________ test statistic)** (ONE TAILED)
- **If H_a: __ > __, *p*-value = $2 \times P(z \geq |$___________ test statistic|)** (TWO TAILED)

Name______________________________

AP Statistics CED 6.5 Daily Video 2 (Skill 4.B)

Interpreting *p*-Value

What Will We Learn?

How do we interpret the *p*-value in a test for a population proportion?

Does Green = More Natural?

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural.

From previous videos:

- H_0: $p = 0.50$ vs H_a: $p > 0.50$, where p = the proportion of all students at this school who would choose the green cup.
- Conditions are met.
- $\hat{p} = 0.60$, $z = 1.10$, and the p-value = 0.1357

Interpreting the *p*-value

In our lemonade study, $\hat{p} = \frac{18}{30} = 0.60$. This is evidence for H_a because $\hat{p} = 0.60 > 0.50$.

The __________ measures how likely it is to get evidence for _______ as __________ as or ______________ than the observed _______________ by chance alone when _______ is true.

So, we would say something like:

"Assuming ______ of all students ____________________ would choose the green cup, there is a _______ probability of getting a ______________ proportion of ______________________ by chance alone in a _____________ sample of ______ students from ____________________."

Favorite Sport

A newspaper reports that 40% of the adults in the U.S. would say football is their favorite sport. The mayor of a town wonders if the proportions of adult in the town who would say football is their favorite sport differs from the national proportion. To Investigate, the mayor selected a random sample of 100 adults in the town and found that 29 named football as their favorite sport. Interpret the *p*-value of 0.0244.

From previous videos:

- H_0: $p = 0.40$ vs H_a: $p \neq 0.40$, where p = the proportion of all adults in the town who would say football is their favorite sport.
- Conditions are met.
- $\hat{p} = 0.29$, $z = -2.25$, and the p-value = 0.0244.

Name_______________________________

"Assuming ______ of ______________________________ would say _______________________ ____________________, there is a _______ probability of getting a ______________ proportion as ___________________ or _____________________ than ________ in _____________ direction by chance alone in a _____________ sample of ______ adults ____________________________.

What Should We Take Away?

How do we interpret the *p*-value in a test for a population proportion?

- **The ___________ measures how likely it is to get evidence for H_a as _______________ or __________________ than the observed evidence by ________________ alone when H_0 is ______________.**
- **Make sure the answer is ___________________.**
- **Don't forget to reference that the _____________ calculation is done with the assumption that __________ is _________.**
- **Remember that the *p*-value _________________ the probability of getting a value as _______________________ or ______________________________ the observed value.**

Name________________________________

AP Statistics CED 6.6 Daily Video 1

Concluding a Test for a Population Proportion

What Will We Learn?

How do we make a conclusion in a test for a population proportion?

Some Preliminary Information…

Small __________ indicate that the observed value of the test statistic would be ____________ if the ___________ hypothesis were __________, and so provide _________________ statistical for the _____________ hypothesis. The smaller the ________, the more ___________________ the statistical evidence for the _________________ hypothesis.

How "__________" does a ____________ need to be to provide statistical evidence for H_a? The __________________ level ∝ is a _____________________ boundary value that we use to determine if a ___________ is _____________ or _____________________. Common significance levels are ∝ = 0.05; ∝ = 0.01, and ∝ =0.10.

Some Preliminary Information….

If the _________ ≤ ∝, we _______ H_0.

In this case, there _____ convincing statistical evidence for H_a.

If the __________ > ∝, we ______________ H_0.

In this case, there __________ convincing statistical evidence for H_a.

Does Green = More Natural?

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural. Is there convincing statistical evidence at the ∝ = 0.05 significance level that more than half of all students at this school would choose the green cup?

From previous videos:

- H_0: $p = 0.50$ vs H_a: $p > 0.50$, where p = the proportion of all students at this school who would choose the green cup.
- Conditions are met.
- $\hat{p} = 0.60$, $z = 1.10$, and the p-value = 0.1357

Here is our conclusion:

"Because the _____________ of 0.1357 ______ ∝ = 0.05, we ______________ H_0. There is not _____________ statistical evidence that _________________ half (0.50) of _________ students at ___________________ would choose the green cup."

Name________________________________

Cautions About Conclusions

- Conclusions must be justified by ______________ comparing the ______________ to the significance level (∝). If ____ significance level is stated in the item, you should pick a significance level to compare to (e.g., ∝ = 0.05).
- Conclusions must be about the ______________________ hypothesis.
- Conclusions must be in ___________________.
- When you fail to reject the H_0, don't conclude that the H_0 is ____________! This is called "_________________ the null hypothesis." Don't do it!
- When you reject H_0, don't say that you have "_____________" that H_a is true.

Favorite Sport

A newspaper reports that 40% of the adults in the U.S. would say football is their favorite sport. The mayor of a town wonders if the proportions of adult in the town who would say football is their favorite sport differs from the national proportion. To Investigate, the mayor selected a random sample of 100 adults in the town and found that 29 named football as their favorite sport. Do these data provide convincing statistical evidence that the proportion of all adults in the town who would say that football is their favorite sport differs from 0.40? Use ∝ = .10.

Favorite Sport

From previous videos:

- H_0: p= 0.40 vs H_a: p ≠ 0.40, where p = the proportion of all adults in the town who would say football is their favorite sport.
- Conditions are met.
- $\hat{p}$ = 0.29, z = -2.25, and the p-value = 0.0244.

"Because the p-value of ______________________________, we reject ________. There is convincing statistical evidence that the proportion of ___________________________ who would say that __________________ is their ________________________________ differs from _____________."

What Should We Take Away?

How do we make a conclusion in a test for a population proportion?

There are two possible conclusions:

- **Because the p-value of _______________________, we _______________ H_0. There _____ convincing ____________________________ that [state H_a in context].**
- **Because the p-value of _______________________, we _________________ H_0. There ____________ convincing _______________________ that [state H_a in context].**

Remember to:

- **_____________ compare the _______ to _______.**
- **Include a conclusion about H_a, _______________________.**

Name________________________________

AP Statistics CED 6.6 Daily Video 2

Concluding a Test for a Population Proportion

What Will We Learn?

How do we perform a complete significance test for a population proportion?

2005 #4

Some boxes of a certain brand of breakfast cereal include a voucher for a free video rental inside the box. The company that makes the cereal claims that a voucher can be found in 20 percent of the boxes. However, based on their experiences eating this cereal at home, a group of students believes that the proportion of boxes with vouchers is less that 0.2. This group of students purchased 65 boxes of the cereal to investigate the company's claim. The student found a total of 11 vouchers for free video rentals in the 65 boxes.

Suppose it is reasonable to assume that the 65 boxes purchased by the students are a random sample of all boxes of this cereal. Based on this sample, is there support for the students' belief that the proportion of boxes with vouchers is less than 0.2? Provide statistical evidence to support your answer.

2005 #4

1. State the Hypotheses:

   H_0: ___________ vs H_a: ___________, where p = ____.

2. Identify the Significance Level: Since none stated, use ∝ = _________

3. Identify the procedure for the test: ____________________________________

4. Check the conditions:
   - ____________ sample of boxes ☐
   - Assume the sample size (65) ___________ of the ____________________ size ☐
   - ______________________ and ____________________ ☐

2005 #4

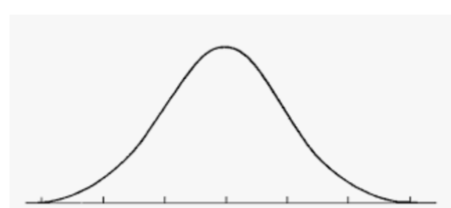
5. Find $\hat{p}$ = ________________ = ____________ (Since this is < 0.2, this is evidence for H_a.)

6. Calculate the test statistic:

   *Note this value is less than one standard deviation below the mean, which does not appear to be that unusual.

$$z = \frac{\underline{\quad} - \underline{\quad}}{\sqrt{\frac{\underline{\quad}(1-\underline{\quad})}{\underline{\quad}}}} = \underline{\qquad\qquad}$$

7. Do Normal Distribution Calculation: Because H_a: $p < .2$, we want to find $P(z \leq$ _________$)$. Draw and shade the graph to the right.



Then use technology or Table A to calculate the actual p-value = ___________________

Name________________________________

2005 #4

8. Make conclusion:

Because the __________________________________, we __________________________. There is _______ convincing statistical evidence that the proportion of _____________________________ ________________________ is less than _______.

What Should We Take Away?

How do we perform a complete significance test for a population proportion?

Make sure to:

- **State the ________ and _________________ hypotheses, making sure to define the ____________________.**
- **Identify the ________________________________.**
- **Identify the _____________________ you are using.**
- **Verify the ________________ for the procedure are ______ (with evidence!)**
- **Calculate the ________________________ and ________________.**
- **Make a conclusion based on the _______________. (You do __________ need to interpret the ______________ unless specifically asked.)**

Name________________________________

AP Statistics CED U6.7 Daily Video 1 (Skill 1.B)

Potential Errors When Performing Tests

What Will We Learn?

How do we identify Type I and Type II errors?

How do we interpret Type I and Type II errors?

Type I and Type II Errors

There are two possible conclusions we can make in a significance test:

Conclusion 1: If the ________________, we reject _____. If this case, there is ___________________ statistical evidence for _________.

If we find convincing evidence for H_a, but H_0 is actually ______, then we made a ___________________. A Type I error occurs when the ________ hypothesis is true and is ___________ (______________________________).

Conclusion 2: If the ________________, we ________ to reject _______. In this case, there is ________ convincing statistical ________________ for H_a.

If we don't find convincing evidence for H_a but H_0 is actually ______, then we made a ______________________. A Type II error occurs when the ________ Hypothesis is false and is ______ __________________ (_______________________________).

Type I and Type II Errors (Fill in the table as you watch the video.)

Table of Errors

DECISION		Actual Population Value	
		H_0 true	H_A true
	Reject H_0		
	Fail to Reject H_0		

Does Green = More Natural?

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each subject was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade. Of the 30 students, 18 stated that the lemonade in the green cup tasted more natural. Is there convincing statistical evidence at the $\alpha = 0.05$ significance level that more than half of all students at this school would choose the green cup?

Name______________________________

From previous videos:

- H_0: p = 0.50 vs H_a: p > 0.50, where p = the proportion of all students at this school who would choose the green cup.

Does Green = More Natural?

Type I: The researchers __________ convincing evidence that ______________________ of all students at this school would choose the green cup, but the actual percentage is __________.

Type II: The researcher ___________________ convincing evidence that more than 50% of all students at this school would choose the green cup, but the actual percent is ______________.

Does Green = More Natural?

Which error is more consequential?

If these student researchers get hired in a marketing department:

- A ___________ would lead them to use green in the branding of a food and beverage items when people ________________________ associate green with being more natural. This seems _______ risk, unless green ink costs a lot.
- A ___________ would lead them to _______ use green in the branding of food and beverage items when people _____________ associate green with being more natural. This would result in a _______________________________________.
- ______________ is more consequential because a ___________ isn't very harmful, but a ______________ would reduce income.

*Note: It is _________ always the case that a _______________ is more consequential, it depends on the _______________ of the problem.

What Should We Take Away?

How do we identify Type I and Type II errors?

A _____________ occurs when the _______ hypothesis is _______ and is ____________________.
This is sometimes called a _______________________.
A _____________ occurs when the _______ hypothesis is _______ and is ____________________.
This is sometimes called a _______________________.

How do we interpret Type I and Type II errors?

Type I: We ________ convincing evidence that [H_a in context], when H_a isn't actually ________.
Type II: We _______________ convincing evidence that [H_a in context], when H_a is actually __________.

Name________________________________

AP Statistics CED U6.7 Daily Video 2 (Skill)

Concluding a Test for a Population Proportion

What Will We Learn?

How do we calculate the probabilities of Type I and Type II errors?

What factors affect the power of a test and the probabilities of errors in significance testing?

Type I Errors

When the ________ hypothesis is ________, the probability of a ________________ is equal to the ______________________________. That is *P*(Type I error) = ∝.

Type I Errors

Because the significance level of ∝ is the probability of a ________________, the consequences of a Type I error ____________ influence our choice of __________________________________. Why not make ∝ = 0.000001 to minimize the ________________ of a Type I error? If ∝ is that small, it makes it ________________ difficult to find convincing evidence for H_a.

All other things being equal, ________________ the probability of a Type I error ________________ the probability of a ________________.

Power and Type II Errors

In many contexts, a ____________ is more consequential.

The probability of _______________ a Type II error = _____________.

The _______________________ is the probability that test will _______________ reject a false null hypothesis. That is, the probability that we ________ find convincing evidence for _________, when _________ is really true.

P(Type II error) = ___________________

Does Green = More Natural?

To investigate if high school students associate the color green with being more natural, two student researchers randomly selected 30 students from their school. Each student was asked to taste two cups of lemonade in random order and state which one tasted more natural. One of the cups was green and the other cup was white. Unknown to the taste testers was the fact that both cups contained the same brand of lemonade.

From previous videos:

- H_0: $p = 0.50$ vs H_a: $p > 0.50$, where p = the proportion of all students at this school who would choose the green cup.

Name______________________________

Does Green = More Natural?

Suppose the researchers use ∝ = 0.05 and that the power of the test against H_a of p = 0.64 is 0.45. Interpret the power in context. Then find the probability of a Type I and a Type II error.

If the _______ proportion of students who would choose the green cup is _______, there is a __________ probability of finding convincing evidence that more that _________ of students would choose the green cup.

P(Type I error) = ∝ = _______

P(Type II error) = 1 – power = _____________ = _______

What Factors Affect Power?

Assuming everything else remains the same, the power of a test will be ________________ (and the probability of a Type II error will be __________________) when:

1. The ___________ size ________.
2. The __________________ level increase (∝) increases.
3. The __________________ decreases.
4. The _________ parameter value is father from the ________.

What Should We Take Away?

How do we calculate the probabilities of Type I and Type II errors?

P(Type I error) = ________

P(Type II error) = ____________

The _____________________ is the probability that at test will ___________ reject a false ___________________________.

What factors affect the power of a test and the probabilities of errors in significance testing?

1. The _________________________ (n)

2. The _________________________ (∝)

3. The _________________________

4. The _____________ between the __________ parameter and the ____________.

Name________________________________

AP Statistics CED U6.8 Daily Video 1 (Skill 4.C)

Confidence Intervals for the Difference of Two Proportions

What Will We Learn?

How do we identify an appropriate confidence interval procedure for a difference in proportions?
How do we verify the conditions for calculating a confidence interval for a difference in proportions?

Please Trees, Don't Leave!

A disease is killing many trees in your state. Random samples of trees from two different large forests, one at high elevation and on at low elevation, reveal that 36 of 240 trees at high elevation and 25 of 200 trees at low elevation died from the disease. Calculate and interpret a 90% confidence interval for the difference (*high-low*) in the proportions of all trees that have died from the disease at these elevations.

Identifying the Procedure

When the goal is to estimate a difference in proportions, we use a ________________________
____.

This procedure is appropriate for estimating a difference in proportions when:

- ______________________ have been selected (one from each of ________ populations),
- Subjects are ______________ assigned to _______ groups in an ____________________.

Check the Conditions

Remember that for ________ inference procedures in AP Statistics, you ___________ verify that the ___________________ for using that procedure are __________.

In general, you should check for:

- _________________ in the methods used to _________________________________, and
- that the appropriate __________________ distribution has the correct ______________.

Conditions: Random Samples

____________________ from _________________________________:

To check for independence:

1. Data are collected using ____________________ samples (one from ________ population).
2. When sampling __________________ replacement, the sample size is less than or equal to _______ of the population size for ______________________________________.

To check that the shape of the sampling distribution is approximately normal:

3. The number of successes and failures in the samples are all at least ____________.
   In other words: $n_1 \hat{p}_1 \geq$ ___, $n_1(1 - \hat{p}_1)$ ___ 10; **AND** $n_2\hat{p}_2 \geq$ ___, $n_2\hat{p}_2$ ___ 10.

Conditions: Treatment Groups

________________ from a randomized ______________________:

To check for independence:

1. Data are collected from _________ groups that have been _________________ assigned in an _______________________.

To check that the shape of the ____________________ distribution is approximately normal:

2. The number of successes and failures in both groups are at least ________.
   In other words: $n_1 \hat{p}_1 \geq$ ___, $n_1(1 - \hat{p}_1)$ ___ 10; **AND** $n_2\hat{p}_2 \geq$ ___, $n_2(1 - \hat{p}_2)$ ___ 10.

Name______________________________

Checking the conditions

Random samples of trees from two different large forests, one at high elevation and on at low elevation, reveal that 36 of 240 trees at high elevation and 25 of 200 trees at low elevation died from the disease. Check if the conditions for calculating a confidence interval are met.

Let ____ = high elevation and ____ = low elevation.

1. A _____________ sample of __________ trees at high elevation was selected. ☐
   A _____________ sample of __________ trees at low elevation was selected. ☐
2. 240 trees is _____________________ of all trees in a ____________ forest. ☐
   200 trees is _____________________ of all trees in a ____________ forest. ☐
3. $n_1\hat{p}_1$ = _____________________ = 36___ 10; $n_1(1 - \hat{p}_1)$ = ___________________ = 204 ____ 10.
   $n_2\hat{p}_2$ = _____________________ = 25___ 10; $n_2(1 - \hat{p}_2)$ = ___________________ = 175 ____ 10.

The __________________ are ___________.

Practice is a Drag

Drag suits are specially designed swimsuits intended to increase drag and make it harder for swimmers to swim. If drag is really increased, swimmers should swim more slowly. Swimmers wear drag suits during practice to increase the intensity of their workouts. Do drag suits really work? Two student researchers randomly assigned 23 swimmers to wear a drag suit and 24 swimmers to wear their regular suits during a 100-meter freestyle race in practice. The recorded whether the swimmers were slower than their average time. Of the 23 swimmers who wore a drag suit, 13 had a time slower than their average. In the group that wore their regular suits, 8 were slower than their average time. Check if the conditions for calculating a confidence interval are met.

Practice is a Drag

Two groups of swimmers were ________________ assigned to wear ____________ suits or _________________ suits. _______________ who wore drag suits were slower than their average, compared to _______________ who wore regular suits.

Let ____ = drag suit and ____ = regular suit.

1. Swimmers were ______________ assigned to wear either drag suits or regular suits. ☐
2. $n_D\hat{p}_D$ = ______________ = ___________ ☐and $n_D(1- \hat{p}_D)$ = ___________________________ ☐
   $n_R\hat{p}_R$ = ______________ = ____________ ☐and $n_R(1 - \hat{p}_R)$ = ___________________________ ☐

What Should We Take Away?

How do we identify an appropriate confidence interval procedure for a difference in proportions?

A _________________________________ for a ____

How do we verify the conditions for calculating a confidence interval for a difference in proportions?

1. _______________________ samples or _______________________ from a randomized experiment.

2. When sampling without replacement, samples sizes are ___________________________________ of the population size. (Doesn't apply to _______________________.)

3. The counts of successes and failures are ______________________________.

$n_1\hat{p}_1 \geq$ ___, $n_1(1 - \hat{p}_1)$ ___ 10; **AND** $n_2\hat{p}_2 \geq$ ___, $n_2\hat{p}_2$ ___ 10

Name________________________________

AP Statistics CED U6.8 Daily Video 2 (Skill 3.D)

Confidence Intervals for the Difference of Two Proportions

What Will We Learn?

How do we determine the margin of error when estimating a difference in proportions?
How do we calculate a confidence interval for a difference in proportions?

Please Trees, Don't Leave!

A disease is killing many trees in your state. Random samples of trees from two different large forests, one at high elevation and on at low elevation, reveal that 36 of 240 trees at high elevation and 25 of 200 trees at low elevation died from the disease. Calculate and interpret a 90% confidence interval for the difference (*high-low*) in the proportions of all trees that have died from the disease at these elevations.

Calculating the Margin of Error

In AP Statistics, confidence intervals have the form:

CI = ____________________ ± _________________________

The margin of error described how much a value of a ________________________________ is likely to vary from the value of the corresponding _____________________________________.

The margin of error is determined by ____________ factors:

- How much the _____________ typically varies from the _____________________.
- How __________________ we want to be in our estimate.

margin of error = (_________________)(_____________________________)

Calculating the Margin of Error

Margin of error = (critical value)(__________________________________)

The standard error of a statistic is an _________________ of the __________________________ of the __________________distribution of the ______________________.

Because we don't know the value of p_1 or p_2, we replace them with $\hat{p}_1$ and $\hat{p}_2$ to get the ____.

______ $= \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Calculating the Margin of Error

margin of error = (______________________)(standard error of statistics)

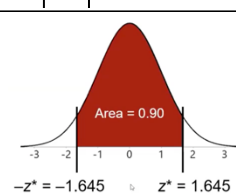
The _______________________ is a multiplier that makes the margin of error _______________ enough to give a ___________ amount of confidence that the ____________________________ the value being estimated.

For confidence intervals for a _____________________________________, the critical values represent the _____________________ encompassing the _______________ C% of the standard normal distribution where C% is the approximate ______________________________ for a proportion.

Calculating the Margin of Error

margin of error = (critical value)(standard error of statistic)

In the diseased trees example, we are asked to construct a 90% confidence interval. To find the _________________ of ________ for a 90% confidence interval, find the boundaries encompassing the ______________ 90% of the _________________ normal curve.



The critical value is z^* = _______

Name________________________________

Calculating the Confidence Interval

CI = ___________ $\pm$ __________________________, which translates to

CI = ______________ $\pm$ (________________)(___________________________)

Define identifiers: Let 1 = ___________________ and Let 2 = ___________________

Calculate $\hat{p}_1$ = _____________________ and $\hat{p}_2$ = ______________________

Identify point estimate = __________________ and z^* = ________________________

Insert values and calculate the CI:

$$CI = (\_\_\_\_ - \_\_\_\_) \pm \_\_\_\_\sqrt{\frac{\_\_(\_\_\_\_\_\_)}{\_\_\_\_\_} + \frac{\_\_(\_\_\_\_\_\_)}{\_\_\_\_\_}}$$

CI = ______________________

Calculating the Confidence Interval

You can find the information for creating a confidence interval for two proportions on your AP Statistics formula sheet. Take a minute to locate and highlight them now.

Who's a Good Dog?

A company that manufactures tick repellent for dogs has developed a new formula that has less odor than their old formula. However, if the new formula doesn't prevent ticks as well as the old formula, dog owners won't buy the new formula. What is the difference in the effectiveness of these two formulas? An experiment was conducted with 160 dogs volunteered by their owners. The dogs were randomly assigned a formula: 80 to the new and 80 to the old formula.

After being treated, the dogs played for 1 hour in a heavily wooded area known to have many ticks. Of the dogs treated with the new formula, 12 got ticks, compared to 25 dogs treated with the old formula. Calculate a 95% confidence interval for the difference (*new – old*) in the true proportion of dogs like the ones in this study that would get ticks.

Who's a Good Dog?

Define identifiers: _______________________ and _________________________

Calculate $\hat{p}_N$ = ________________________ and $\hat{p}_O$ = _________________________

Identify point estimate = __________________ and z^* = ________________________

Insert the values and calculate CI:

$$CI = (\_\_\_\_ - \_\_\_\_) \pm \_\_\_\_\sqrt{\frac{\_\_(\_\_\_\_\_\_)}{\_\_\_\_\_} + \frac{\_\_(\_\_\_\_\_\_)}{\_\_\_\_\_}}$$

The CI = _____________________

What Should We Take Away?

How do we determine the margin of error when estimating a difference in proportions?

margin of error = (_____________________**)(**_________________________________**)**

_______________ = $z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

How do we calculate a confidence interval for a difference in proportions?

CI = _________________ $\pm$ _________________________________

______________ = $(\hat{p}_1 - \hat{p}_2) \pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Name________________________________

Name______________________________

AP Statistics CED U6.9 Daily Video 1 (Skill 4.B)

Claims: Confidence Intervals for Population Proportion Differences

What Will We Learn?

How do we interpret a confidence interval for a difference of proportions?

How do we justify a claim based on a confidence interval for a difference in proportions?

Please Trees, Don't Leave!

A disease is killing many trees in your state. Random samples of trees from two different large forests, one at high elevation and on at low elevation, reveal that 36 of 240 trees at high elevation and 25 of 200 trees at low elevation died from the disease. Calculate and interpret a 90% confidence interval for the difference (*high-low*) in the proportions of all trees that have died from the disease at these elevations.

Interpreting the Confidence Interval

In general, here is how to interpret a confidence interval for a population parameter:

"We are _____ confident that the ______________ from ____ to ____ captures the __________ __________________."

From a previous video (6.8, Video 2), the 90% confidence interval for the diseased trees example is -0.029 to 0.079. So, we would interpret that as:

" We are ______ confidence that the interval from _________________ captures the __________________________ in the proportion of ______________________ that have died from the disease."

Justifying a Claim

Wildlife biologists must decide how to use their limited resources to battle the disease. If the disease seems to be more damaging at one of the elevations, they will use more resources to fight the disease at that elevation. Is there convincing evidence that the disease is more lethal at one of the elevations?

Ponder: If the disease is _____________ lethal in both forests, the __________________ in the proportion of all trees that have died from the disease would be _______.

Because ________ is in the 90% confidence interval (_____________________), 0 is a _________________ value for $p_H - p_L$, the difference (*high- low*) in the proportion of ________ trees that have died from the disease. Thus, there ______________ convincing evidence that the disease is more lethal at one of the elevations.

Name______________________________

Who's a Good Dog?

A company that manufactures tick repellent for dogs has developed a new formula that has less odor than their old formula. An experiment was conducted with 160 dogs volunteered by their owners. The dogs were randomly assigned a formula: 80 to the new and 80 to the old formula. Of the dogs treated with the new formula, 12 got ticks, compared to 25 of 80 dogs treated with the old formula.

Who's a Good Dog?

The 95% confidence interval for the difference (*new-old*) in the true proportion of dogs such as the ones in the study that would get ticks is -0.1625 ± 0.1282.

(a) Interpret the confidence interval.

(b) Based on the interval, is there convincing evidence that the new formula is better than the old formula at preventing ticks on dogs such as the ones in this study?

Who's a Good Dog?

(a) Interpret the confidence interval.

Create the interval: -0.1625 ± 0.1282 = ________________________

Write interpretation: We are ___________ confidence that the interval from _______ to _______ captures the ___________________ in the __________ proportion of dogs such as theses that would get ticks when using these two repellents.

Who's a Good Dog?

(b) Based on the interval, is there convincing evidence that the new formula is better than the old formula at preventing ticks on dogs such as the ones in this study?

Because __________________ in the interval (_____________________) are less than _____, there _______ convincing evidence that the _________ formula is better than the _________ formula at preventing ticks on dogs _____________________________.

What Should We Take Away?

How do we interpret a confidence interval for a difference of proportions?

"We are _______ confident that the interval from _______ to _______ captures the _________ ____________________________."

How do we justify a claim based on a confidence interval for a difference in proportions?

- **If ____ of the values in the confidence interval are ___________________ with the claim, there ______ convincing evidence for the claim.**
- **If ___________________ of the values in the confidence interval are ________________ with the claim, there ___________ convincing evidence for the claim.**

Name______________________________

AP Statistics CED U6.9 Daily Video 2 (Skill 4.D)

Claims: Confidence Intervals for Population Proportion Differences

What Will We Learn?

How do we construct and interpret a confidence interval for difference in proportions?

2006 B #2

A large company has two shifts – a day shift and a night shift. Parts produced by the two shifts must meet the same specifications. The manager of the company believes that there is a difference in the proportions of parts produced with specifications by the two shifts. To investigate this belief, random samples of parts that were produced on each of these shifts were selected. For the day shift, 188 of its 200 selected parts met specifications. For the night shift, 180 of its 200 selected parts met specifications.

(a) Use a 96 percent confidence interval to estimate the difference in the proportions of parts produced within specifications by the two shifts.

(b) Based only on this confidence interval, do you think that the difference in the proportions of parts produced within specifications by the two shifts is significantly different from 0? Justify your answer.

2006 B #2, part (a)

(a) Use a 96 percent confidence interval to estimate the difference in the proportions of parts produced within specifications by the two shifts.

State: ____ (*day-night*) in the proportions of _______ parts produced within specifications by the two shifts.

Label Identifiers: Let 1 = _____________ and Let 2 = ______________

Identify Procedure: ____

Check conditions:

1. Random: ____ ☐

____ ☐

2. Since sampling _____________ replacement must check the 10% Condition:

____ ☐

____ ☐

3. Large Counts Condition:

__________________________ ☐ __________________________ ☐

__________________________ ☐ __________________________ ☐

The conditions are _______________.

2006 B #2

Calculate $\hat{p}_1$ = ______________________ and $\hat{p}_2$ = ______________________

Calculate critical value (z^*) by finding the middle 96% of area in a standard normal distribution = _______________

Create confidence interval by plugging in values.

$$CI = (\_\_\_ - \_\_\_) \pm \_\_\_\sqrt{\frac{\_\_(\_\_\_\_)}{\_\_\_} + \frac{\_\_(\_\_\_\_)}{\_\_\_}}$$

The confidence interval is from:

____

*Note: verify your calculation using technology.

Name______________________________

2006 B #2 Interpret the confidence interval in context: We are ___________________ that the interval from _______________________ capture the difference __________________ in the proportions of ________ parts produced within specifications for these two shifts.
2006 B #2 (b) Based only on this confidence interval, do you think that the difference in the proportions of parts produced within specifications by the two shifts is significantly different from 0? Justify your answer. Because ______ is in the _________ confidence interval, ____ is a plausible value for the _________________ (*day-night*) in the proportions of _____ parts produced within specifications by the two shifts. There ___________ convincing _________________ that the proportions of parts produced within specifications by the two shifts is ___________________ different from _____.
2006 B #2, BONUS FEATURE! Interpret the confidence level. (This wasn't asked on the exam!) If _____ possible ___________ samples of _______ part from each shift were selected and a confidence interval was constructed from each pair of samples, then __________ of all these intervals would succeed in ________________ the difference (*day-night*) in the proportion of all parts produced within specifications by the two shifts.
What Should We Take Away? How do we construct and interpret a confidence interval for difference in proportions? **Be sure to:** • **______________ the difference in proportions you are trying to estimate.** **Indicate the _____________ of the difference.** **Define any ______________ you use.** • **______________ the _________________ you are using.** • **______________ that the _______________ for the procedure are met.** • **______________ the confidence interval.** • **______________ the interval in ______________.**

Name______________________________

AP Statistics CED U6.10 Daily Video 1 (Skill 1.F)

Setting up a Test for the Difference of Two Population Proportions

What Will We Learn?

How do we state a null hypothesis in a test for a difference in proportions?

How do we state an alternative hypothesis in a test for a difference in proportions?

The Structure of Unit 6

Unit 6 focuses on _________________ data that can be summarized by calculating the proportion of _____________ in a sample or treatment group.

Topics 6.2 – 6.7 → confidence intervals and significance tests for a ___________________________.

Topics 6.8 – 6.9 → confidence intervals for a ____.

Topics 6.10 – 6.11 → significance tests for a ____.

Some Reminders…

Confidence Intervals

- used to ____________ the value of a ____.
- an interval of ____________ values for a parameter based on the ____________ data.

Significance Tests

- used to ____________________ about the value of a population _________________.
- assess whether the evidence supporting a claim is likely or unlikely to happen by ________________________.

Effective Eyedrops?

Bacterial conjunctivitis (pink eye) causes eye irritation in those unlucky enough to contract it. Can eye drops containing azithromycin make the recovery faster? In one clinical trial, eyedrops containing azithromycin were compared to a placebo eyedrop among 279 patients suffering from pink eye. The results? Of the 130 patients who were randomly assigned azithromycin drops, 82 were cured within a week. Of the 149 patients who were randomly assigned placebo drops, 74 were cured within a week. Do these results give convincing evidence that azithromycin is more effective at curing pink eye than a placebo for subjects like the ones in this study?

Null Hypothesis

In a statistical test, the ________ hypothesis is often a claim of "_____ difference" or "_____ change."

In the effective eyedrops example, the null hypothesis is that there is ___________________ ________________________like the ones in the study who would be ____________________ ___________ using either of the drops.

Symbolically the null hypothesis would say:

H_0: _____ = _____ OR H_0: ___________ = ____

where p_1 = the _________ proportion of pink eye patients like the ones in _____________ who would be cured by azithromycin drops

p_2 = the _________ proportions of pink eye patients like the one in ______________ who would be cured by placebo drops

Alternative Hypothesis

In a statistical test, the ___________________ hypothesis is the claim that we _________ to support with ________________ from the data collected.

Name______________________________

In the effective eyedrops example, the researchers wanted to know if the azithromycin drops were ________ effective. So, the _________________ hypothesis is that __________________ of patients ____________________________ who would be cured by azithromycin drops is ________________ than the proportion for placebo drops.

Symbolically the hypotheses would say:

H_0: ______ = ______ OR H_0: ___________ = ____

H_A: ______ = ______ OR H_A: ___________ = ____

where p_1 = the _________ proportion of pink eye patients like the ones in ______________ who would be cured by azithromycin drops

p_2 = the _________ proportions of pink eye patients like the one in _______________ who would be cured by placebo drops

Stating Hypotheses: Summary

For hypotheses about a ___________________ of proportions:

The ________ is a statement about _________, typically H_0: $p_1 - p_2$ = _____

The __________________ always contains a strict _________________, typically

H_A: $p_1 - p_2$ ____ 0, H_A: $p_1 - p_2$ ____ 0, OR H_A: $p_1 - p_2$ ____ 0

- When the inequality is _____ or _____, the alternative is called "________________"
- When the inequality is __________, the alternative is called "two-sided."
- The choice of alternative is determined by the ______________________________ and should be stated __________________________________ begins.
- Never refer to statistics (such as _____ or ______) in the ________________!
- Remember to ___________ an parameters you use. (Specifically define ____ and ____)

A Bright Idea

A marketing agent wonders if there is a difference in the percentage of residents in two cities who have purchase sunglasses during the last 12 months. Of 400 randomly selected residents of Soltown, 314 purchased sunglasses in the last 12 months. Meanwhile, a random sample of Brightville residents revealed that 452 of 550 residents purchased sunglasses in the last 12 months. Do these data give convincing evidence of a difference in the proportions of residents who have purchased sunglasses in the past 12 months in these two cities? State appropriate hypotheses for the agent's test.

*Note: When you are asked if "data gives convincing evidence" about a claim that is an indication that you are looking at a conducting a significance test.

A Bright Idea

Do these data give convincing evidence of a difference in the proportions of residents who have purchased sunglasses in the past 12 months in these two cities?

H_0: __________________ OR H_0: ___________________

H_A: __________________ OR H_A: ___________________

where: p_1 = ____

p_2 = ____

Pay close attention to the wording when choosing the appropriate inequality!!

What Should We Take Away?

How do we state a null hypothesis in a test for a difference in proportions?

H_0: $p_1 = p_2$; remember to clearly ________________ parameter

How do we state an alternative hypothesis in a test for a difference in proportions?

H_A: $p_1 > p_2$ H_A: $p_1 < p_2$ H_A: $p_1 \neq p_2$

STATS MEDIC

Name________________________________

AP Statistics CED 6.10 Daily Video 2 (Skill)

Concluding a Test for a Population Proportion

What Will We Learn?

How do we identify an appropriate significance test procedure for a difference in proportions?
How do we verify the conditions for performing a significance test for a difference of proportions?

Effective Eyedrops?

Bacterial conjunctivitis (pink eye) causes eye irritation in those unlucky enough to contract it. Can eye drops containing azithromycin make the recovery faster? In one clinical trial, eyedrops containing azithromycin were compared to a placebo eyedrop among 279 patients suffering from pink eye.

The results? Of the 130 patients who were randomly assigned azithromycin drops, 82 were cured within a week. Of the 149 patients who were randomly assigned placebo drops, 74 were cured within a week. Do these results give convincing evidence that azithromycin is more effective at curing pink eye than a placebo for subjects like the ones in this study?

Effective Eyedrops?

Bacterial conjunctivitis (pink eye) causes irritation in those unlucky enough to contract it. Can eye drops containing azithromycin make the recovery faster?

From a previous video:

H_0: $p_1 = p_2$ OR H_0: $p_1 - p_2 = 0$
H_A: $p_1 = p_2$ OR H_A: $p_1 - p_2 = 0$

where p_1 = the ________ proportion of pink eye patients __________________________ who would be cured by azithromycin.

p_2 = the _________ proportion of pink eye patients _________________________ who would be cured by azithromycin.

Identifying the Procedure

You have already learned about a significance test for a population proportion involving _________ sample. But in this case, there are __________ groups.

When the goal is to test a claim about a _________________ in proportions, we use a:
____.

This procedure is appropriate when:

- _________ random samples have been selected (________ from each of _______ populations, or
- subjects are randomly assigned to _________ groups in an experiment.

Checking the Conditions

Remember that for all inference procedures in AP Statistics you must ____________ that the ________________ for using that procedure are met.

In general, you should check for:

- _____________________ in the methods used to collect the data, and
- that the appropriate ___________________ distribution has the correct shape.

In the case of a ________________________ test for a difference of proportions, the conditions for ___________ random samples look a little different than for _________ groups in a __________________ experiment.

Name________________________________

Checking the Conditions

Before checking the for a two-sample z test for a difference in proportions, first compute the ____ in both samples or groups. For the effective eyedrops example, the __________ hypothesis is the treatments are __________ effective. If that's true, then the best estimate of their effectiveness is to ____________ all the patients from __________________ into ______________________ and calculate the proportion who were cured.

Checking the Conditions

Azithromycin group: _____ cured out of ________.

Placebo group: ______ cured out of ________.

Combined (______________) proportion of successes $\hat{p}_c$:

$$\hat{p}_c = \frac{\text{combined number of successes}}{\text{combined number of observations}} = \frac{X_1+X_2}{n_1+n_2} = \frac{\_\_\_ + \_\_\_}{\_\_\_ + \_\_\_} = \frac{\_\_\_}{\_\_\_} = \_\_\_\_\_\_\_\_ = \text{(pooled proportion)}$$

Conditions: Random Samples

For a ____________________ for a ______________ in proportions involving ____________ samples from _________ populations:

To check for Independence:

1. Data are collected using _________________________ samples. (one for each population.
2. When sampling ___________ replacement, the sample size is less than or equal to _________ of the population size for ___________ samples.

To check that the _________ of the ______________ distribution is approximately normal:

3. The ___________ number of successes and failure in the samples are all at least ____.
   In other words, $n_1\hat{p}_c$ ____ 10, $n_1(1 - \hat{p}_c)$ ____ 10, $n_2\hat{p}_c$ ____ 10, and $n_2(1 - \hat{p}_c)$ ____ 10

Conditions: Experiments

For a ____ involving _______ groups from a randomized __________________:

To check for independence:

1. Data are collected from ________ groups that have been _______________ assigned in an experiment.

To check that the shape of the ____________________ distribution is approximately normal:

2. The _____________ number of successes and failures in ___________ groups are all at least ________.
   In other words, $n_1\hat{p}_c$ ____ 10, $n_1(1 - \hat{p}_c)$ ____ 10, $n_2\hat{p}_c$ ____ 10, and $n_2(1 - \hat{p}_c)$ ____ 10

Checking the Conditions:

In the effective eyedrops study, 82 of the 130 patients that used azithromycin drops and 74 of the 149 patients given placebo drops were cured after one week. Check if the conditions for performing the significance test are met.

Let 1 = ______________ and 2 = ________________

1. A group of 130 patients were _______________ assigned to use azithromycin drops. ☐
   A group of 149 patients were _______________ assigned to use placebo drops. ☐
2. $n_1\hat{p}_c$ = ____________ = _______ ____10, and $n_1(1 - \hat{p}_c)$ = _______________ = _______ ____ 10 ☐
   $n_2\hat{p}_c$ = ____________ = _______ ____10, and $n_2(1 - \hat{p}_c)$ = _______________ = _______ ____ 10 ☐

Name______________________________

A Bright Idea

A marketing agent wonders if there is a difference in the percentage of residents in two cities who have purchase sunglasses during the last 12 months. Of 400 randomly selected residents of Soltown, 314 purchased sunglasses in the last 12 months. Meanwhile, a random sample of Brightville residents revealed that 452 of 550 residents purchased sunglasses in the last 12 months. Do these data give convincing evidence of a difference in the proportions of residents who have purchased sunglasses in the past 12 months in these two cities? Check if the conditions for performing a significance test are met.

A Bright Idea

Check if the conditions for performing a significance test are met.

First, calculate the ___________ (pooled) proportion of ___________: $\hat{p}_c = \dfrac{\_\_\_\_ + \_\_\_\_}{\_\_\_\_ + \_\_\_\_} =$ ___________

Let ____ = Soltown and ____ = Brightville

1. A ____________ sample of _______ Soltown residents was selected.
   A ____________ sample of _______ Brightville residents was selected.
2. 400 residents is __________________ of the population in a city. ☐
   550 residents is __________________ of the population in a city. ☐
3. $n_1\hat{p}_c =$ _____________ = _______ ____10, and $n_1(1 - \hat{p}_c) =$ ________________ = _______ ____ 10 ☐
   $n_2\hat{p}_c =$ _____________ = _______ ____10, and $n_2(1 - \hat{p}_c) =$ ________________ = _______ ____ 10 ☐

*Note: Remember to use the combined (pooled) proportion in both of these.

What Should We Take Away?

How do we identify an appropriate significance test procedure for a difference in proportions?

A ____

How do we verify the conditions for performing a significance test for a difference of proportions?

1. **Two _________________ _________________ samples or two _____________ from a randomized ______________________.**
2. **When sampling ________________ replacement, sample sizes are _______________________ to _________ of the population sizes (This DOESN'T apply to _______________________.)**
3. **The ________________ counts of successes and failures are _____ at least ___________. In other words, $n_1\hat{p}_c$ ____ 10, $n_1(1 - \hat{p}_c)$ ____ 10, $n_2\hat{p}_c$ ____ 10, and $n_2(1 - \hat{p}_c)$ ____ 10**

Name________________________________

AP Statistics CED 6.11 Daily Video 1 (Skill 3.E)

Carrying Out a Test for the Difference of Two Population Proportions

What Will We Learn?

How do we calculate an appropriate test statistic in a test for a difference in proportions?

How do we calculate a *p*-value in a test for a difference of proportions?

Effective Eyedrops?

Bacterial conjunctivitis (pink eye) causes eye irritation in those unlucky enough to contract it. Can eye drops containing azithromycin make the recovery faster? In one clinical trial, eyedrops containing azithromycin were compared to a placebo eyedrop among 279 patients suffering from pink eye. The results? Of the 130 patients who were randomly assigned azithromycin drops, 82 were cured within a week. Of the 149 patients who were randomly assigned placebo drops, 74 were cured within a week. Do these results give convincing evidence that azithromycin is more effective at curing pink eye than a placebo for subjects like the ones in this study?

Effective Eyedrops?

From a previous video:

H_0: $p_1 = p_2$ OR H_0: $p_1 - p_2 = 0$

H_A: $p_1 = p_2$ OR H_A: $p_1 - p_2 = 0$

where p_1 = the ________ proportion of pink eye patients ________________________________ who would be cured by azithromycin.

p_2 = the _________ proportion of pink eye patients _______________________________ who would be cured by azithromycin.

Calculating a Test Statistic

In the effective eyedrops study: $\hat{p}_1 - \hat{p}_2 = \frac{\_\_\_}{\_\_\_} - \frac{\_\_\_}{\_\_\_} =$ ______________

This is ___________ for H_a: ___________ > 0 because $\hat{p}_1 - \hat{p}_2$ = ________ > 0

We want to know how ____________ it is to get evidence for H_a _______________ or _______________ by ________________ alone when H_0 is _______________.

After verifying the _______________ are met, calculate the _________________________ test statistic:

standardized test statistics = $\frac{\_\_\_\_\_\_\_\_ - \_\_\_\_\_\_\_\_}{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_}$

Calculating a Test Statistic

For a ___________________ for the difference of proportions the standardized test statistic is:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_c(1-\hat{p}_c)}{n_1} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}_c(1-\hat{p}_c)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

The second version is the one you will find on the AP Statistics Formula Sheet.

where $\hat{p}_c$ is the combined (__________________) proportion of successes.

$$\hat{p}_c = \frac{\text{combined number of successes}}{\text{combined number of observations}} = \frac{X_1 + X_2}{n_1 + n_2} = \text{(pooled proportion)}$$

Name______________________________

Calculating a Test Statistic

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1-\hat{p}_C)}{n_1} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}_C(1-\hat{p}_C)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

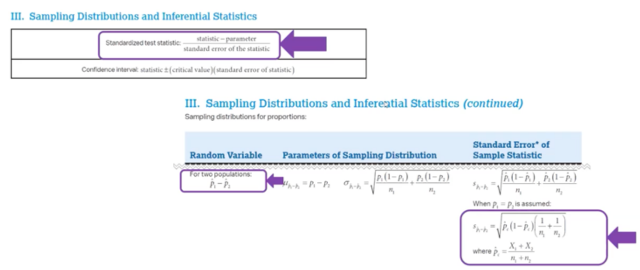
First, calculate the combined (____________) proportion: $\hat{p}_C = \frac{\_\_\_ + \_\_\_}{\_\_\_ + \_\_\_} = \frac{\_\_\_}{\_\_\_} = \_\_\_\_\_\_\_\_$

Then plug in values to find the test statistic:

$$z = \frac{\left(\frac{\_\_\_}{\_\_\_} - \frac{\_\_\_}{\_\_\_}\right) - 0}{\sqrt{\_\_\_(1 - \_\_\_)\left(\frac{1}{\_\_\_}+\frac{1}{\_\_\_}\right)}} = \_\_\_\_\_\_\_\_$$

Calculating a Test Statistic

All of the pieces for this calculation can be found on your AP Statistics Formula Sheet.

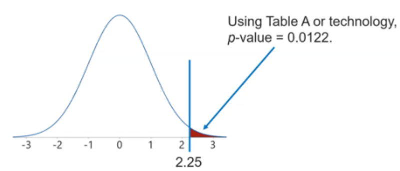


Calculating a *p*-value

Once we have calculated the standardized test statistic, use the ______________ normal distribution to calculate the *p*-value. The _______________ is the probability of observing a test statistics as _______________ or _________ extreme than the observed test statistics when the ___________ hypothesis and _____________________ are assumed to ______________.

Calculating a *p*-value

Because our alternative hypothesis is H_a: $p_1 - p_2 >$ _______, we want to find $P(z \geq 2.25)$.



A Bright Idea

A marketing agent wonders if there is a difference in the percentage of residents in two cities who have purchase sunglasses during the last 12 months. Of 400 randomly selected residents of Soltown, 314 purchased sunglasses in the last 12 months. Meanwhile, a random sample of Brightville residents revealed that 452 of 550 residents purchased sunglasses in the last 12 months. Do these data give convincing evidence of a difference in the proportions of residents who have purchased sunglasses in the past 12 months in these two cities? Do these data give convincing evidence of a difference in the proportions of residents who have purchase sunglasses in the pas 12 months? Calculate the standardized test statistic and the *p*-value.

A Bright Idea

From a previous video:

H_0: $p_S = p_B$ or H_0: $p_S - p_B = 0$

H_a: $p_S \neq p_B$ or H_a: $p_S - p_B \neq 0$

where p_S = the proportion of ______ Soltown residents who purchased sunglasses in the past 12 months.

P_B = the proportion of ______ Brightville residents who purchased sunglasses in the past 12 months.

STATS MEDIC

Name________________________________

Calculate a Test Statistic

First, calculate the two sample statistics: $\hat{p}_S$ = ________________ and $\hat{p}_B$ = _________________

Then, calculate the combined (pooled) proportion:

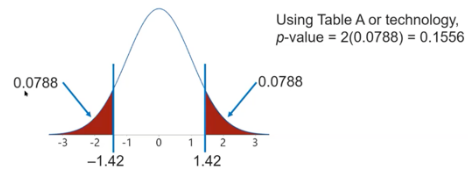
$$\hat{p}_c = \frac{\_\_\_\_ + \_\_\_\_}{\_\_\_\_ + \_\_\_\_} = \frac{\_\_\_\_}{\_\_\_\_} = \_\_\_\_\_\_\_\_\_\_$$

Finally, calculate the z statistic:

$$z = \frac{\left(\frac{\_\_\_\_}{\_\_\_\_} - \frac{\_\_\_\_}{\_\_\_\_}\right) - 0}{\sqrt{\_\_\_\_(1 - \_\_\_\_)\left(\frac{1}{\_\_\_\_} + \frac{1}{\_\_\_\_}\right)}} = \_\_\_\_\_\_\_\_\_\_$$

Calculating a *p*-value

Because our alternative hypothesis is H_a: $p_S - p_B \neq 0$, we want to find $P(z \leq -1.42) + P(z \geq 1.42) \Longrightarrow 2 \times P(z \geq |-1.42|)$



p-value = ____________

What Should We Take Away?

How do we calculate an appropriate test statistic in a test for a difference in proportions?

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1-\hat{p}_C)}{n_1} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}_C(1-\hat{p}_C)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

How do we calculate a *p*-value in a test for a difference of proportions?

- **If H_a: $p_1 > p_2$ OR H_a: $p_1 - p_2 > 0$, *p-value* = $P(z \geq$ ______________ test statistic)**
- **If H_a: $p_1 < p_2$ OR H_a: $p_1 - p_2 < 0$, *p*-value = $P(z \leq$ ______________ test statistic)**
- **If H_a: $p_1 \neq p_2$ OR H_a: $p_1 - p_2 \neq 0$, *p*-value = $2 \times P(z \geq |$___________ test statistic|)**

Name________________________________

AP Statistics CED 6.11 Daily Video 2 (Skill 4.B)

Carrying Out a Test for the Difference of Two Population Proportions

What Will We Learn?

How do we interpret the *p*-value for a significance test for a difference in proportions?

How do we state a conclusion for a significance test for a difference in proportions?

Effective Eyedrops?

Bacterial conjunctivitis (pink eye) causes eye irritation in those unlucky enough to contract it. Can eye drops containing azithromycin make the recovery faster? In one clinical trial, eyedrops containing azithromycin were compared to a placebo eyedrop among 279 patients suffering from pink eye. The results? Of the 130 patients who were randomly assigned azithromycin drops, 82 were cured within a week. Of the 149 patients who were randomly assigned placebo drops, 74 were cured within a week. Do these results give convincing evidence that azithromycin is more effective at curing pink eye than a placebo for subjects like the ones in this study?

Effective Eyedrops?

From a previous video:

H_0: $p_1 = p_2$ OR H_0: $p_1 – p_2 = 0$ AND H_A: $p_1 = p_2$ OR H_A: $p_1 – p_2 = 0$

where p_1 = the ________ proportion of pink eye patients ______________________________ who would be cured by azithromycin.

p_2 = the _________ proportion of pink eye patients _____________________________ who would be cured by azithromycin.

Additionally, we know that conditions are met.

Interpreting a *p*-value

From previous videos: $\hat{p}_1$ - $\hat{p}_2$ = _______, z = __________, and *p*-value = _________________

$\hat{p}_1$ - $\hat{p}_2$ = _______, is evidence ________ H_a: $\hat{p}_1$ - $\hat{p}_2$ >0 because $\hat{p}_1$ - $\hat{p}_2$ = _______ > 0.

The *p*-value measures how _______________ it is to get evidence for H_a as ______________ or more _______________ than the _______________ evidence by ___________ alone when H_0 is ____________.

Interpreting a *p*-value

"Assuming H_0 is _________, there is a __________ probability of getting a _________________ in proportions of ________________________ or ______________________________, by chance alone in the _________________ assignment (_______________ samples)."

Interpret the *p*-value of the effective eyedrops example:

Assuming the difference (_______________ - _______________) in the ________ proportion of pink eye patience ________________________ in this experiment who ____________ be cured is 0, there is a _______________ probability of getting a __________________ in proportions of __________ or _________________, by chance alone in the ________________ assignment.

Stating a Conclusion

___________ *p*-values ⟹ test statistic is _______________ to occur by ________________ chance alone.

___________ *p*-values ⟹ test statistic is _______________ to occur by ________________ chance alone.

- Because the *p*-value of ___________ ≤ ∝ = _______, we reject H_0.
  There is ________________ statistical evidence that [state H_a ________________].
- Because the *p*-value of ___________ > ∝ = _______, we fail to reject H_0.
  There is ____________________ statistical evidence that [state H_a ________________].

STATS MEDIC

Name________________________________

Stating a Conclusion

No significance level was stated in the effective eyedrops example, so we'll use ∝ = _______, which is the most ______________ significance level.

- Because the p-value of ___________≤ ∝ = _______, we reject H_0.
  There is ________________ statistical evidence that the ________________ (azithromycin – placebo) in the proportions of patients ____ who _____________ be cured within a week is _______________________.

A Bright Idea

A marketing agent wonders if there is a difference in the percentage of residents in two cities who have purchase sunglasses during the last 12 months. Of 400 randomly selected residents of Soltown, 314 purchased sunglasses in the last 12 months. Meanwhile, a random sample of Brightville residents revealed that 452 of 550 residents purchased sunglasses in the last 12 months.
The p-value for a two-sample z test of the difference (*Soltown – Brightville*) in proportions is 0.01556. Interpret the p-value and make a conclusion at the ∝ = 0.10 significance level.

A Bright Idea: Interpret *p*-value

From a previous video:
H_0: $p_S = p_B$ or H_0: $p_S - p_B = 0$
H_a: $p_S \neq p_B$ or H_a: $p_S - p_B \neq 0$ $\hat{p}_S - \hat{p}_B$ = _____________, z = ______________, and p-value = _________

Interpret the p-value:
Assuming the _________________ (*Soltown – Brightville*) in the proportions of _________ residents who have purchased sunglasses during the last 12 months is ___, there is a __________ probability of getting a __________________ in sample proportions of _________ or one ______________ different in _______________ direction, by _____________ alone in the ________________ samples.

A Bright Idea: State Conclusion

Because the p-value of _________ > ∝ = ________, we _______________ H_0. There is __________ convincing statistical __________________ that the difference (*Soltown – Brightville*) in the proportions of ___________ residents who ______________________ sunglasses during the last 12 months is _________ equal to _______.

What Should We Take Away?

How do we interpret the p-value for a significance test for a difference in proportions?
"Assuming H_0 is _________, there is a __________ probability of getting a ________________ in proportions of ________________________ or _______________________________, by chance alone in the _________________ assignment (_______________ samples)."

How do we state a conclusion for a significance test for a difference in proportions?

- **Because the p-value of ___________≤ ∝ = _______, we reject H_0.**
  There is ________________ statistical evidence that [state H_a _______________].
- **Because the p-value of ___________ > ∝ = _______, we fail to reject H_0.**
  There is ____________________ statistical evidence that [state H_a _______________].

Name________________________________

AP Statistics CED 6.11 Daily Video 3 (Skill 4.E)

Carrying Out a Test for the Difference of Two Population Proportions

What Will We Learn?

How do we perform a complete significance test for a difference in proportions?

2012 #4

A survey organization conducted telephone interviews in December 2008 in which 1,009 randomly selected adults in the United States responded to the following question.

> At the present time, do you think television commercials are an effective way to promote a new product?

Of the 1,009 adults surveyed, 676 responded "yes." In December 2007, 622 of 1,020 randomly selected adults in the United States had responded "yes" to the same question. Do the data provide convincing evidence that the proportion of adults in the United States who would respond "yes" to the question changed from December 2007 to December 2008?

Hypotheses

Null: ______________________________________

Alternative:________________________________

This is a ______ tailed test!

where: p_1 = the proportion of _______ U.S. adults who ________________ respond "yes" to the question in ______________________________.

p_2 = the proportion of _______ U.S. adults who ________________ respond "yes" to the question in ______________________________.

No __________________________ was stated, so we'll use ____________________.

Identify the Procedure and Check the Conditions (Be sure to put a ✓ by your conditions!)

____

1. **Independence:** Two _______________________________ samples of U.S. adults, __________ from December 2008 and __________ from December 2007.

2. **10% Condition:**

_________________ is less than ________ of ______ U.S. adults in December 2008

_________________ is less than ________ of ______ U.S. adults in December 2007

Combined _____________ number of successes: $\hat{p}_c = \frac{__ + __}{__ + __} = \frac{__}{__} = $ __________

3. **Expected Counts ≥ 10:**

____

____

The conditions __________________ met

Name______________________________

Calculating the Test Statistic and *p*-value

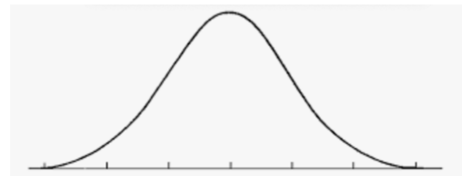
$\hat{p}_1$ = _____________ $\hat{p}_2$ = _____________ $\hat{p}_C$ = _____________

Test Statistic:

$$z = \frac{(\_\_\_\_ - \_\_\_\_) - 0}{\sqrt{\_\_\_\_\_(1 - \_\_\_\_\_)\left(\frac{1}{\_\_\_\_} + \frac{1}{\_\_\_\_}\right)}} = \_\_\_\_\_\_\_\_$$

Calculate *p*-value:
(Sketch and lable the standard normal curve).
Calculate the *p*-value using Table A or techonology.

The *p*-value = _______________



Hint: Is this a one or two-tailed test?

Conclusion and Interpretation (Use the interpretation your teacher suggests.)

Because the *p*-value of ____________________________________, we would ___________ H_0.
There __________________________ statistical evidence that ___________________________
____
____.

What Should We Take Away?

How do we perform a complete significance test for a difference in proportions?

Make sure to:

- **State the _________ and ______________ hypotheses, define any _________________, and indicate the ______________ in the difference (>, < or ≠)**
- **Identify the __________________________, (if none provided use ______________)**
- **Identify the __________________________ you are using.**
- **Verify that the _________________ for the procedure are met (with ______________!)**
- **Calculate the _______________________ and ______________.**
- **Make a _____________________ based on the p-value. (You do *not* need to interpret the *p*-value unless specifically asked.)**