Learning Goals:

- Given a claim about a population proportion, determine null and alternative hypotheses.
- Recognize the logic behind a hypothesis test and how it relates to the P-value.
    - Compare P-values to a level of significance to draw a conclusion
    - State conclusions to hypothesis tests using the language of "reject the null" or "fail to reject the null" and statistical significance

Introduction:

In the last activity we focused on constructing confidence intervals. In this activity we will focus on hypothesis testing. Confidence intervals and hypothesis testing are the two types of statistical inference we will study this semester.

We use a hypothesis test when we want to find a "yes" or "no" answer to a claim about a population parameter.

1) Which of the following questions can be answered by a hypothesis test?

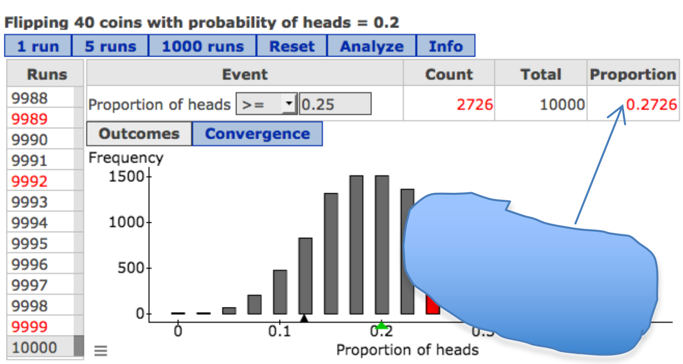
- What is the average amount of money that community college students receive in financial aid?
- Is the proportion of CA community college students that qualify for the BOG fee waiver greater than 40%?
- Do the majority of community college students qualify for federal student loans?

Now we will walk through an example of a hypothesis test. The hypothesis test is the same thinking we did in Activity 18.4, so we will review a problem from that activity and add to it the vocabulary and notation of a hypothesis test.

Steps in a hypothesis test:
Step 1: Determine the hypotheses.
Step 2: Collect the data and report the sample results.
Step 3: Assess the data.
Step 4: State a conclusion

Example: According to a 2013 survey of college students conducted by Citi and Seventeen Magazine, 20% of college students have a credit card when they start college. Is the percentage higher for community college students this year?

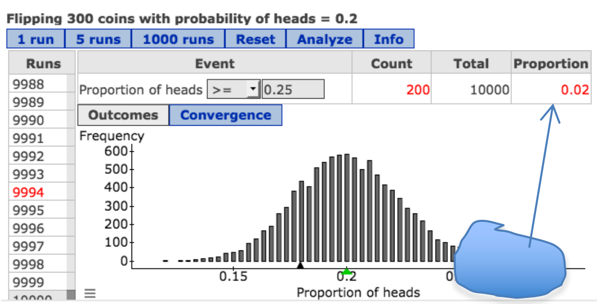
In a survey of a random sample of 40 community college freshmen, 25% have a credit card.

What we did before	**Notation and vocabulary of a hypothesis test**
We start with the hypothesis that 20% of community college freshmen have a credit card this year. Our claim is that the percentage is higher.	*Step 1: Determine the hypotheses.* Ho: p=0.20 Ha: p>0.20 p=proportion of community college freshmen with a credit card this year
To test the hypothesis, we select a random sample of community college freshmen. Suppose that the data shows that 25% of the sample has a credit card.	*Step 2: Collect the data and report the sample results.* $\hat{p}$ = 0.25
Obviously, 25% is greater than 20%. But we need to determine if this 5 percentage point difference is typical or unusual when we look at samples coming from the population. Here is the distribution of 10,000 sample proportions from samples of 40 community college freshmen.  Visually, we can see that 0.25 is not unusual in this distribution so the associated probability is large. About 27% of the time, we expect a random sample to have 25% or more with a credit card when 20% of the population owns one.	*Step 3: Assess the data.* Find the P-value. The P-value is 0.27, which is large. This indicates that the survey result of 25% is not unusual when sampling from a population where Ho is true.
Are we right that more than 20% of the all community college freshmen own a credit card? In other words, is our claim true? Even though more than 20% of the sample had a credit card, *the difference is not large enough* to support our claim that the population proportion is also larger than 20%. Therefore, we conclude that our *claim is not supported by the data.*	*Step 4: State a conclusion* The sample evidence is not statistically significant. The observed difference between 0.25 and 0.20 can be attributed to sampling variability. We do not have enough evidence to conclude that more than 20% of the population of community college freshmen owns a credit card. Fail to reject Ho.

You may be wondering … how does the conclusion change when the P-value is small?

In 18.4 we also ran a simulation of 10,000 samples of 300 students for this scenario. Let's *assess the data* and *state a conclusion* with these larger samples.

<table>
<tr>
<td>Visually, we can see that 0.25 is unusual in this distribution so the associated probability is small.



About 2% of the time, we expect a random sample to have 25% or more with a credit card when 20% of the population owns one.</td>
<td>*Step 3: Assess the data.*

Find the P-value.

The P-value is 0.02, which is small.

This indicates that the survey result of 25% is unusual when sampling from a population where Ho is true.</td>
</tr>
<tr>
<td>Are we right that more than 20% of the all community college freshmen own a credit card? In other words, is our claim true?

It is surprising to see random samples with 25% or more owning a credit card when 20% of the population owns one. In other words, this *difference is large enough* to support our claim that the <u>population proportion</u> is also larger than 20%.

Therefore, we conclude that our *claim is true.*</td>
<td>*Step 4: State a conclusion*

The sample evidence is statistically significant. The observed difference between 0.25 and 0.20 cannot be attributed to sampling variability.

We have enough evidence to conclude that more than 20% of the population of community college freshmen owns a credit card.

Reject Ho in favor of Ha.</td>
</tr>
</table>

You may be wondering … how small does a P-value have to be in order to accept the claim as true and reject the null hypothesis?

Since the hypothesis test asks us to answer "Is the claim true?" (yes or no), it is common practice to agree ahead of time on the definition of "unusual" sample results so that we can answer this question without any haggling. This definition of "unusual" is called a **significance level.**

If the P-value is small enough to be less than the significance level, then the sample data is considered unusual. We conclude the sample data is statistically significant. This is strong evidence that our claim is true. (Reject Ho in favor of Ha.)

If the P-value is larger than the significance level, then the sample data is not considered unusual. We conclude the sample data is not statistically significant. We do not have sufficient evidence to support our claim. (Fail to reject Ho.)

Group work

2) In 2001 polls indicated that 74% of Americans favored mandatory testing in public schools as a way to rate the school. Your local school board conducts a survey to determine if a smaller proportion of residents in the district now support the use of mandatory test results in rating schools. They find that 68% of the residents are in support this year.

   a) Which one of the following pairs of hypotheses fit this scenario?

   - Ho: $p=0.74$, Ha: $p<0.68$
   - Ho: $p=0.68$, Ha: $p<0.68$
   - Ho: $p=0.74$, Ha: $p<0.74$
   - Ho: $p=0.68$, Ha: $p<0.74$

   b) What does p represent in the hypotheses? Write a sentence that clearly defines p.

3) Forty-six percent of high school students ages 12 to 17 in the United States have had sexual intercourse, according to a 2014 study by the Official Journal of the American Academy of Pediatrics (OJAAP).

   Is this percentage higher in inner city high schools?

   Suppose that a study finds that 52% of inner city high school students have had sexual intercourse. The sample data is used to test the hypotheses: Ho: $p=0.46$, Ha: $p>0.46$. The P-value is 0.02.

a) Which conclusion is supported by this P-value?

- The sample evidence is unusual when sampling from a population with $p=0.46$, so the data is statistically significant.

- The sample evidence is not unusual when sampling from a population with $p=0.46$, so the data is not statistically significant.

b) Which conclusion is supported by this P-value?

- Reject Ho in favor of Ha

- Fail to reject Ho

c) What can we conclude? (Choose all that apply)

- The difference between 0.46 and 0.52 is statistically significant.

- This survey provides enough evidence to conclude that the percentage of all high school students engaging in sexual intercourse is higher in the inner cities.

- This study does not suggest that inner city high school students are engaging in sexual intercourse at higher rates.

- The difference between 0.46 and 0.52 can be explained by expected sampling variability.

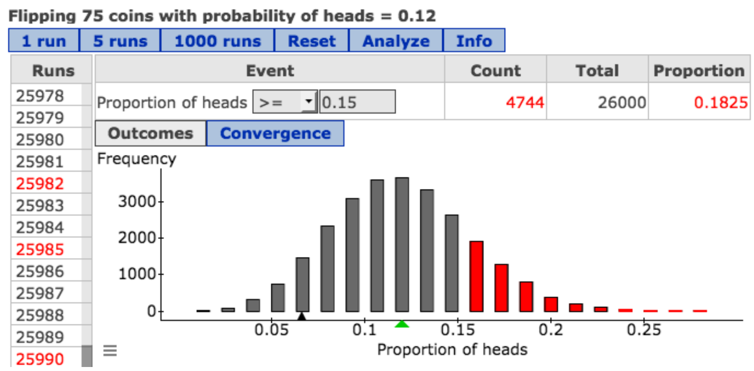
4) A 2011 California Student Survey found that in any 30-day period, almost 12 percent of students in the 9th grade admit to using drugs at least once at school.
Source: http://www.huffingtonpost.com/marsha-rosenbaum/drug-education_b_3906983.html

To investigate this issue in your community, suppose that the school board hires a firm to conduct a survey of a random sample of 9th graders to determine if drug usage is greater than 12% at schools in the district. Seventy-five students respond to the anonymous survey and 11 admit to drug usage at school. (11 out of 75 is about 15%)

The sample data is used to test the hypotheses:

Ho: $p=0.12$ (12% of the population of 9th graders have used drugs at school)
Ha: $p>0.12$ (More than 12% of the population of 9th graders have used drugs at school)

Here is the StatCrunch image from the simulation of selecting 26,000 random samples of 75 students.



a) What is the P-value?

b) Which conclusion is supported by this P-value?

- The sample evidence is unusual when sampling from a population with p=0.12, so the data is statistically significant.

- The sample evidence is not unusual when sampling from a population with p=0.12, so the data is not statistically significant.

c) Which conclusion is supported by this P-value?

- Reject Ho in favor of Ha

- Fail to reject Ho

d) What can we conclude? (Choose all that apply)

- The difference between 0.12 and 0.15 is statistically significant.

- This survey provides enough evidence to conclude that the percentage of 9th graders using drugs at school is higher than 12%.

- This study does not suggest that the drug usage rate for 9th graders in this district is higher than 12%.

- The difference between 0.12 and 0.15 can be explained by expected sampling variability.

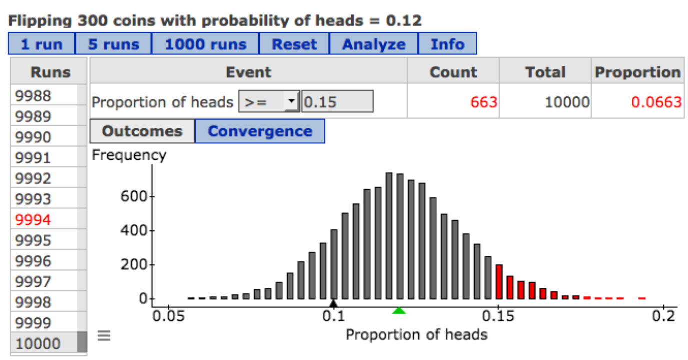
5) The school board is happy to hear that drug usage by 9th graders in the district is not higher than the statewide rate. But a concerned citizen raises a concern about the relatively small sample size. The school board agrees to redo the study with a larger sample.

Because of the controversy over the first study, the school board sets a 5% significance level to alleviate disagreements over the conclusions from the second study.

This time the research firm is able to get 300 9th graders to respond; 45 admit to drug use at school, which is again 15%. The sample data is again used to test the hypotheses:

Ho: p=0.12 (12% of the population of 9th graders have used drugs at school)
Ha: p>0.12 (More than 12% of the population of 9th graders have used drugs at school)

Here is the StatCrunch image from the simulation of selecting 10,000 random samples of 300 students.



a) What is the P-value?

b) Considering the significance level, which conclusion is supported by this P-value?

- The sample evidence is unusual when sampling from a population with p=0.12, so the data is statistically significant.

- The sample evidence is not unusual when sampling from a population with p=0.12, so the data is not statistically significant.

c) Considering the significance level, which conclusion is supported by this P-value?

- Reject Ho in favor of Ha

- Fail to reject Ho

d) What can we conclude? (Choose all that apply)

- The difference between 0.12 and 0.15 is statistically significant.
- This survey provides enough evidence to conclude that the percentage of 9th graders using drugs at school is higher than 12%.
- This study does not suggest that the drug usage rate for 9th graders in this district is higher than 12%.
- The difference between 0.12 and 0.15 can be explained by expected sampling variability.

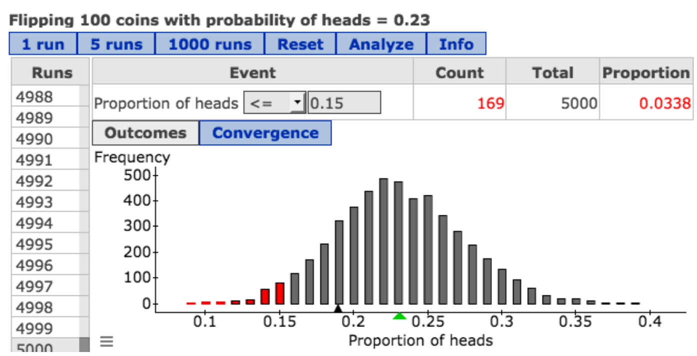
6) In an article titled "Tattoos Becoming More Accepted at Work", CBS News reported in 2007 that 23% of college students were tattooed. Let's use this as a hypothesis for the proportion the population of LMC students who are tattooed. Test this hypothesis against a level of significance of 5%.

Suppose that in a random sample of 100 LMC students, 15% are tattooed. What can we conclude?

a) State the hypotheses (use Ho and Ha notation)

b) What is the sample proportion that we will use to test the hypotheses?

c) Determine the P-value based on the simulation.



d) State a conclusion

7) What questions do you have at this point about hypothesis testing?