Learning Goals:

- Under appropriate conditions, conduct a hypothesis test about a population mean. State a conclusion in context.
- Interpret the P-value as a probability.

Introduction:

Up to this point, we have had quite a bit of experience conducting hypothesis tests. But we have always worked with categorical data and tested hypotheses about proportions. In this activity we learn to conduct a hypothesis test about a population mean. We will see that the process and the logic of the hypothesis test are very similar to what we did previously.

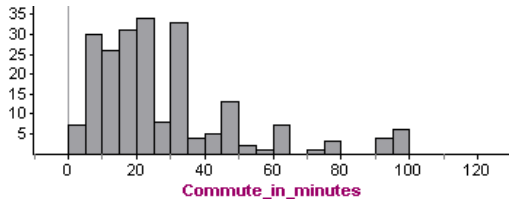
Example:

According to a 2015 study, the average American worker spends about 25 minutes a day commuting to (or from) work. Using U.S. Census data, we will test the hypothesis that commute times are no different now than in the year 2000.

Source: http://www.cheatsheet.com/business/5-cities-with-the-most-brutal-commutes.html/?a=viewall

State the hypotheses:

Collect the data:
A random sample of 215 adults is selected from the 2000 Census. The mean commute time is 25.7 minutes with a standard deviation of 21.78 minutes. Here is a histogram of the data.



- The sample mean of 25.7-minutes from the year 2000 is obviously very close to the 25-minute commute time in 2015. Why do we need to do a hypothesis test? Can't we just say that there is no statistically significant difference in the commute time in the years 2000 and 2015?

Assess the data:

- The skew in this data suggests that the distribution of commute times in the population in the year 2000 is also skewed.

  Verify that we can use a T-curve to model the distribution of sample means despite the likely skew in the population's commute times.

- Calculate the T-score:

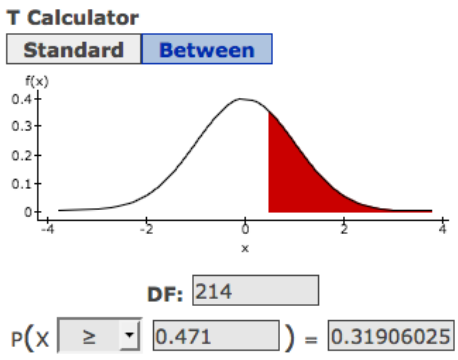
$$T = \frac{statistic - parameter}{standard\ error} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{25.7 - 25}{21.78/\sqrt{215}} = 0.471.$$

- Use the StatCrunch T-calculator to find the P-value. Note that the T-calculator requires you to calculate and enter degrees of freedom and a T-score.

  What is the P-value for this test?

  How does it compare with the StatCrunch print-out?

  Shade the entire P-value in the T-model.



One sample T hypothesis test:
μ : Mean of population
H_0 : μ = 25
H_A : μ ≠ 25

Hypothesis test results:

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	25.7	1.4853837	214	0.47125871	0.6379

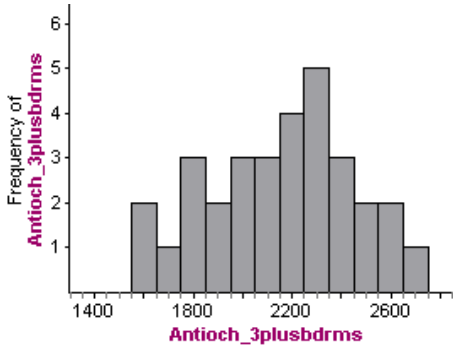
State a conclusion:

Group work

1) For her statistics project, a student wants to determine whether rental rates are lower in Antioch than the county average. According to her research, the average monthly rent paid for a 3+bedroom house in Contra Costa County is $2500.
(Source: http://www.zilpy.com/US/California/Contra_Costa_County/)

She randomly selects thirty 3+bedroom houses in Antioch by randomly selecting 6 rentals from the tulia.com list each week for 5 weeks.

The sample mean is $2,137.52 with a standard deviation of $299.83



a) Explain why she can conduct a hypothesis test using this data.

b) Her hypotheses are Ho:μ=2500, Ha:μ<2500. What does μ represent in terms of rents? Be specific!

c) Here is the StatCrunch output for her hypothesis test. What can she conclude about 3+bedroom rentals in Antioch?

One sample T hypothesis test:
μ : Mean of population
H_0 : μ = 2500
H_A : μ < 2500

Hypothesis test results:

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	2137.52	54.741218	29	-6.6217014	<0.0001

d) What does a standard error of 54.74 mean in this context?

e) What does the T-Stat tell us in this context?

2) For their project, a group of statistics students investigate the features that distinguish a large egg from an extra large egg. With some internet research, they learn that USDA guidelines are based on a carton of 12 eggs. The minimum net weight per dozen for large eggs is set at 24 ounces.

Source: http://www.fsis.usda.gov/wps/portal/fsis/topics/food-safety-education/get-answers/food-safety-fact-sheets/egg-products-preparation/shell-eggs-from-farm-to-table/ct_index

Each team member buys two cartons of large eggs from different grocery stores in the area, for a total of 6 cartons. Because there are so many brands, they choose the third brand from the left on each grocery store shelf.

They assume that weights of dozens of large eggs will be normally distributed and they feel that their process for selecting cartons will produce a random sample. Therefore, they proceed with a T-test.

Here is the StatCrunch printout for their project.

One sample T hypothesis test:
μ : Mean of population
$H_0 : \mu = 24$
$H_A : \mu > 24$

Hypothesis test results:

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	24.25	0.12247449	5	2.0412415	0.0484

a) Which of the following are correct ways for them to state their conclusions?

- At a 5% level of significance, our sample gives strong evidence that the mean weight of a dozen large eggs is greater than 24 ounces.
- At the 5% level of significance, the amount that our sample differs from a mean weight of 24 ounces is statistically significant. Therefore, we reject the null hypothesis in favor of the alternative.
- The 0.25-difference between 24.25 and 24 is too small to be statistically significant at the 5% level. It is not surprising to see random samples vary from 24 ounces as much as our sample.
- We fail to reject the null hypothesis because our sample is only off by 0.25 ounces, which is less than 5%.

b) Which if the following is an accurate interpretation of the P-value?

- If it is true that a random sample of a dozen large eggs weighs 24.25 ounces or more, then there is a 4.8% chance that the mean weight of all cartons of a dozen large eggs is 24 ounces.

- If it is true that all cartons of a dozen large eggs weigh 24 ounces on average, then there is a 4.8% chance that a random carton of a dozen large eggs will weigh 24.25 ounces or more.

- There is a 4.8% chance that the mean weight of all cartons of a dozen large eggs is greater than 24 ounces.

- There is a 4.8% chance that the mean weight of a random sample of a dozen large eggs will weigh 24 ounces.

3) A group of statistics students wants to determine if LMC students work more than 20 hours a week on average. They ask 20 randomly selected students who are working in the math lab on a Tuesday afternoon at 2:15, "Do you work more than 20 hours a week?"

   Will these students be able to run a T-test to determine if LMC students work more than 20 hours a week? Why or why not?

4) A group of statistics students wants to investigate how fast on average they can download a 4-minute song from iTunes using an iPhone and an AT&T phone plan that advertises up to 6Mbps download speed.

   They want to test the hypothesis that the mean download time is greater than 6 seconds. They plan to use a T-test.

   What advice do you have for them about designing their study?

5) The following is an excerpt from the Food and Agriculture Organization of the United Nations (highly summarized).

   In testing meat contamination, the indicator bacteria most commonly used in testing should not exceed 100 per cm for the criterion "good microbiological standard."

   Source: http://www.fao.org/docrep/010/ai407e/AI407E24.htm

   In a T-Test of the following hypotheses,

   Ho: The meat is not spoiled (mean bacteria count is 100 per cm)
   Ha: The meat is spoiled (mean bacteria count is greater than 100 per cm)

   a) Describe Type I and Type II errors for this situation.

   b) Which is the more serious error? Why?

   c) What can be done to avoid making Type I and Type II errors?