

Author: Brenda Gunderson, Ph.D., 2015

License: Unless otherwise noted, this material is made available under the terms of the Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License: <http://creativecommons.org/licenses/by-nc-sa/3.0/>

The University of Michigan Open.Michigan initiative has reviewed this material in accordance with U.S. Copyright Law and have tried to maximize your ability to use, share, and adapt it. The attribution key provides information about how you may share and adapt this material.

Copyright holders of content included in this material should contact open.michigan@umich.edu with any questions, corrections, or clarification regarding the use of content.

For more information about how to attribute these materials visit: <http://open.umich.edu/education/about/terms-of-use>. Some materials are used with permission from the copyright holders. You may need to obtain new permission to use those materials for other uses. This includes all content from:

Attribution Key

For more information see: <http://open.umich.edu/wiki/AttributionPolicy>

Content the copyright holder, author, or law permits you to use, share and adapt:



Creative Commons Attribution-NonCommercial-Share Alike License



Public Domain – Self Dedicated: Works that a copyright holder has dedicated to the public domain.

Make Your Own Assessment

Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright.



Public Domain – Ineligible. Works that are ineligible for copyright protection in the U.S. (17 USC §102(b)) *laws in your jurisdiction may differ.



Content Open.Michigan has used under a Fair Use determination
Fair Use: Use of works that is determined to be Fair consistent with the U.S. Copyright Act (17 USC § 107)
*laws in your jurisdiction may differ.

Our determination DOES NOT mean that all uses of this third-party content are Fair Uses and we DO NOT guarantee that your use of the content is Fair. To use this content you should conduct your own independent analysis to determine whether or not your use will be Fair.

Lab 6: Learning about a Population Mean

Objective: In this lab, you will learn an important statistical technique that will allow you to answer the question, “*Was our observation due to chance, or is it more significant?*” The objective is to guide you through the ideas behind tests of statistical significance and the statistical language involved. This In-Lab Project discusses the one-sample t -test for a population mean, as well as providing practice in answering the question of whether an event can be attributed to chance.

Application: Derek is considering attending the University of Michigan and he wants to consider the cost of rent in Ann Arbor to help make his decision. Ann Arbor’s website claims that the average price for apartments in Ann Arbor is \$800 a month. However, Derek’s friends lead him to believe that the average price for all apartments in Ann Arbor is actually greater than \$800 a month. Derek would like to conduct a hypothesis test to see if his friends’ statement is supported.

Overview:

Hypothesis Tests: A **one-sample t test** is a procedure designed to assess what the evidence provided by the data says about some statement about the population mean. The key elements of the one-sample t test include: a null and alternative hypothesis, assumptions that must be checked, a test statistic t and its p -value, a decision, and a conclusion.

Hypothesis Testing Steps:

1. Determine appropriate null and alternative hypotheses about μ , and set α .
2. Check assumptions for performing the test.
3. Calculate the test statistic and determine the p -value.
4. Evaluate the p -value and report a conclusion in context of the problem.

The first step of a hypothesis test is to identify the hypotheses; this step is crucial, as it dictates the procedures for the remainder of the test. A one-sample t -test is used to test whether the mean of a quantitative variable is significantly different from some value. This value, the **test value** or μ_0 , is given in the null hypothesis ($H_0: \mu = \mu_0$), and is often taken to be zero (i.e., $H_0: \mu = 0$). In Derek’s test, his null hypothesis is $H_0: \mu = 800$.

The **alternative hypothesis**, H_a (or sometimes H_1), represents the experimenter’s new model, or what the experimenter would like to show. The alternative hypothesis can take three different forms – it may be a denial of the null hypothesis (uses \neq ; called a **two-sided test**), or it may specify a direction of interest (uses $>$ or $<$; called a **one-sided test**). Notice that we never test for equality in H_a . Derek’s alternative hypothesis will be $H_a: \mu > 800$.

The test relies on two key assumptions: (1) the data can be considered a **random sample** from the larger population of interest and (2) the data are observations from a **normally distributed population**. The Central Limit Theorem allows the assumption of normality to be more relaxed if our sample size, n , is large. (Generally, “large” means more than 30 observations, although it also depends somewhat on how serious the data depart from normality.) The sampling distribution applet that you worked with demonstrated the CLT using a *large* sample size of 25. Thus we will accept *at least 25* as large enough for inference about means.

The purpose of a significance test is to assess whether or not the observed data are consistent with the null hypothesis (within reasonable bounds of sampling variability). If the data seem unlikely to occur if the null hypothesis is assumed true, then we would reject the null hypothesis statement.

To help us make this decision, we use a **test statistic**, which represents a summary of the data. Going back to the application, Derek samples 40 students at the University by asking about their monthly rent. From the sample he computes a sample mean rent of \$900 as well as a standard error for this sample mean. This test statistic takes the form: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$.

The **test statistic tells us how many standard errors the sample mean, \bar{x} , is from the test value, μ_0** . Derek calculated a test statistic of $t = 1.2$. His interpretation would state: The sample mean rent of \$900 per month is 1.2 standard error away from the hypothesized mean rent of \$800.

In hypothesis testing, another frequently reported value is the **p -value**, a number that is used to indicate the degree of significance of the data. The p -value is the probability of getting a test statistic as extreme or more extreme than the observed value of the test statistic, assuming the null hypothesis is true.

In Derek's case, the p -value is the probability of observing a test statistic of 1.2 or larger, if the population mean rent in Ann Arbor really was 800 dollars more month. Note that more "extreme" means in the direction of H_a . His test statistic, t , has a t -distribution with $n-1$ degrees of freedom (df), where n represents the sample size. In Derek's test, the t statistic has a t (39) distribution.

We must decide in advance how much evidence against H_0 we will require for rejection. This designated amount of evidence is called the level of significance, denoted by α (alpha). Common values of α are 0.01, 0.05, and 0.10. If the p -value is less than or equal to α , Derek would make the decision to reject H_0 . If Derek makes the decision to reject the null hypothesis, the results of the test are said to be statistically significant at level α .

A "significant" result in the statistical sense does not necessarily imply an "important" result in the practical sense. It simply means that such a difference from the null hypothesis is not likely to happen just by chance.

Formula Card

Population Mean	
Parameter	μ
Statistic	\bar{x}
Standard Error	
$s.e.(\bar{x}) = \frac{s}{\sqrt{n}}$	
Confidence Interval	
$\bar{x} \pm t^* s.e.(\bar{x})$	$df = n - 1$
Paired Confidence Interval	
$\bar{d} \pm t^* s.e.(\bar{d})$	$df = n - 1$
One-Sample t-Test	
$t = \frac{\bar{x} - \mu_0}{s.e.(\bar{x})} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$df = n - 1$
Paired t-Test	
$t = \frac{\bar{d} - 0}{s.e.(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}}$	$df = n - 1$

Warm-Up: Check Your Understanding

A researcher wants to determine if a filling machine at a cereal plant fills the boxes well on average or if the machine needs to be replaced. The machine is intended to fill the packages with 28 grams of cereal. He wants to test the machine to see if the average weight of cereal is less than the intended 28 grams. He takes a sample of 30 cereal boxes and measures the weight of each box. The histogram of the weights indicates some skewness.

Which result allows him to conduct the one sample t-test even though his histogram shows some skewness?

He performs a one sample t test and finds a statistically significant p -value. If the significance level was 5%, what are possible options for the p -value?

0.0067 0.043 0.14 0.22

ILP: How Fast is your Stats 250 Lab at the Tangrams game?

Background: The Tangram is a dissection puzzle that is claimed to have originated in China, and players must try to make the shapes out of seven pieces of varying shapes. The most common goal of the game is to make a square out of the seven pieces, but the goal can be changed to a varying number of shapes. In today's lab, you will be completing the "House of Tangrams" puzzle, with your time to complete the puzzle recorded. A previous Stats 250 GSI claims that her lab completed the puzzle in 120 seconds, on average – can your lab beat this?

Task: Perform a test to assess if the average time to complete the "House of Tangrams" puzzle is significantly faster than 120 seconds.

Complete the "House of Tangrams" puzzle.

Your GSI will provide you with a link to access the Tangrams puzzle online. Some tips before you begin:

- Use an alias when playing so you can find the row corresponding to you in the dataset. It does not need to be your real name, but it can be if you'd like.
- Please, only play the game once, as repeated entries will be added to the dataset and will cause our results to be incorrect!
- To rotate a piece clockwise, double-click it (and hold alt to rotate counter-clockwise). If you are having trouble rotating multiple times, it may help to slightly move the cursor after each click.
- If you are attending remotely, please make a quick post in the backchannel saying that you are starting the tangrams puzzle. Reply to that same post when you are done so I can keep track of who still needs to complete the puzzle.

We will need to wait until all the results are entered before we download the data set, so review the overview section of this lab (before the Warm-Up) and when done, you can start working on the next page while you wait for your GSI to announce that the data is ready to download.

Procedure: The appropriate inference procedure for this scenario is the one-sample t test and the specific parameter of interest is _____ .

Hypothesis Test: You can complete the steps for conducting a test of hypotheses.

1. State the Hypotheses: H_0 : _____ = _____ and H_a : _____ ,
where _____ represents:

Determine Alpha: We were told the significance level was 5%.

Remember: Your hypotheses and parameter definition should always be a statement about the **population(s)** under study.

2. Checking the Assumptions

- a. For this scenario, we need to assume that the data are a random sample. To check this assumption, we would make a _____ plot (if there was time order) of the observations, and look for _____ .
- b. We also need to assume that the responses come from a normally distributed _____ . To check this assumption, we would make a _____ plot.
- c. Create the qqplot of the data on time to complete the "House of Tangrams" puzzle and comment.

- d. Based on your sample size, is the assumption of normality for the population of all times to complete the "House of Tangrams" puzzle? Why or why not?

3. Importing Tangrams data into R Commander

When the GSI tells you that everyone has completed the tangrams game, you can then download the data. To do this, click on "Get Game Data" on the page from the link that your GSI gave you earlier. The GSI will give you a group name for your class. Make sure that you select "CSV" as the type to return.

CSV files are "Comma Separated Values," so each item in a row is separated by a column. Most Microsoft Excel spreadsheets have the option to export data as CSV files, so you can use the following importing method to import your own data for analysis into R Commander.

To import this CSV data in R Commander, go to **Data > Import data > From text file**. This gives some options on how to import text -- you can give the dataset a name like "Tangrams" or whatever suits you. For our file, the **Field Separator** will be **commas**, as this is a comma-separated-values file. After that, you can click **OK**, and find the file that you downloaded from the tangrams website.

4. Compute the Test-Statistic and Calculate the p -value

Test-Statistic

- a. Before conducting the test, we will compute some summary statistics used in computing a test statistic. Go to **Statistics > Summaries > Numerical summaries** to generate these, and fill out the table below. Be sure to check the standard error option under the **Statistics** tab.

Summary Statistics			
Mean	Std. Dev (s)	Sample Size (n)	Std. Error

- b. Generate the t -test output using **Statistics > Means > Single-Sample T Test**. Enter a test value of _____ and specify the correct direction for the alternative hypothesis. (This should match your answer in question 1) Use this to fill out the following output table for this test. Note that when conducting a one-sided test, R will provide a confidence *bound* (instead of a confidence interval), which is not taught in this course.

One Sample T Results		
t	df	p -value

- c. What distribution does *the test statistic follow* if the null hypothesis is true?

Note: This will be the model used to find the p -value.
This is not the same as the distribution of *the population that the data were drawn from*.

Calculate the p -value:

- d. Draw a picture of the p -value, with labels for the distribution and x-axis.

4. Evaluate the p -value and Conclusion

Evaluate the p -value:

What is your decision at a 5% significance level? **Reject H_0** **Fail to Reject H_0**

Remember: Reject H_0 \Leftrightarrow Results statistically significant Fail to Reject H_0 \Leftrightarrow Results not statistically significant

Conclusion:

What is your conclusion *in the context of the problem*?

Note: Conclusions should always include a reference to the population parameter of interest. Conclusions should not be too strong; you can say that you have sufficient evidence, but do NOT say that we have proven anything true or false.

5. Providing a Confidence Interval Estimate

Note: In order to get a 95% confidence interval with R Commander, you will need to run a two-sided test (regardless of what test you are conducting).

- a. Generate the two-sided test output and report the 95% confidence interval (include units):

The 95% confidence interval estimate for the population mean time to complete the

“House of Tangrams” puzzle was found to go from _____ to

_____.

- b. How would the 99% confidence interval compare to the 95% interval in part (a)?

Cool-Down: Check Your Understanding

Provide an **interpretation of the test statistic value** that you found in lab to test if the average time to complete the “House of Tangrams” puzzle is significantly faster than 120 seconds.

How would you explain the meaning of the **95% confidence level** used in making the confidence interval for the population mean time to complete the “House of Tangrams” puzzle above?

