

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 4: Hypothesis Testing for a Population Proportion

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 lab first presents a general overview of testing. Then, the In-Lab Project discusses the large sample Z-test for a population proportion, as well as providing practice in answering the question of whether an event can be attributed to chance.

Application: Lucy is an Ann Arbor resident who is considering the possibility of running for a seat on the City Council. To decide if she should invest time and money in her campaign, she wants to conduct some research to see if there is evidence that she could win the election (i.e. she could receive a majority of all the votes).

Overview:

Hypothesis Tests: A **test of hypotheses** (or **significance test**) is a procedure designed to assess what the evidence provided by the data says about some statement about a population parameter. When we conduct this hypothesis test for Lucy, we are interested in testing if the population proportion of votes that would be cast for Lucy is greater than 0.5, ensuring her victory.

Hypothesis Testing Steps:

1. Determine appropriate null and alternative hypotheses 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 the context of the problem.

The first step of this hypothesis test is to identify the hypotheses; this step is crucial, as it dictates the procedures for the remainder of the test. The **null hypothesis**, H_0 , represents the status quo or statement of no effect. The **alternative hypothesis**, H_a (or sometimes H_1), represents the experimenter's new model, or what the experimenter would like to show.

For our application regarding Lucy, we could define p to represent the true proportion of all Ann Arbor residents who would vote for Lucy. We want to test the hypothesis $H_0: p = 0.5$ against the alternative $H_a: p > 0.5$.

Both of the hypotheses are postulated about the same **population** proportion. 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 .

The purpose of a significance test is to assess whether or not the observed data are consistent with the null hypothesis (within the reasonable bounds of sampling variability). If we collect our sample and compute a population proportion of votes for Lucy that is much higher than the 0.5 we specified in the null hypothesis, we have evidence to reject that null hypothesis.

To help us make this decision, we use a **test statistic**, which represents a summary of the data. When we conduct a test about a population proportion, we are performing a Large Sample Z test. The test statistic for this Large sample z test has the following form:

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

This p_0 is the value specified in the null hypothesis. For Lucy's test: $p_0 = 0.5$. The sample proportion of people who will vote for Lucy is our \hat{p} . The test statistic tells us how many standard errors the sample proportion, \hat{p} , is from the test value, p_0 .

The test relies on two key assumptions: that we have a random sample (our sample of responses is representative of the large population of all such responses), and that our sample size is large enough. To achieve a large enough sample size, we need at least 10 hypothesized yes answers and at least 10 hypothesized no answers, or said another way, under the null hypothesis, would we expect to see at least 10 yes answers and at least 10 no answers.

Conditions: $np_0 > 10$ and $n(1 - p_0) > 10$

The Z test statistic has a known probability distribution (under the null hypothesis), and will be examined for evidence in favor of or against H_0 . Under the null hypothesis that the true proportion of votes for Lucy is equal to 0.5, the test statistic Z has a standard normal distribution.

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.* For Lucy's test, the p-value is the probability of getting a z test statistic like we did or greater, assuming that the true proportion of votes for Lucy is equal to 0.5.

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 , we make the decision to reject H_0 . For Lucy's test, if we encounter a p-value smaller than the significance level, we would decide to reject the hypothesis that the true proportion of votes for Lucy is exactly 0.5.

If we 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.

Of course, no procedure is perfect, and as such, there are two types of errors possible during hypothesis testing. If the null hypothesis is true but the decision is to reject H_0 , then we say that a **Type I error** has occurred. If we commit a Type I error for Lucy's test, we would reach the decision that Lucy will win the election by receiving more than 50% of the votes, when in fact the true proportion of votes cast for Lucy is 0.5. Lucy would invest her time and money in her campaign when in fact she will not win the election.

A **Type II error** occurs when the alternative hypothesis is true, but we fail to reject H_0 . A Type II error could be committed in Lucy's test if we decide Lucy will lose the election by receiving 50% of the votes or less, when in fact Lucy will receive greater than 50% of the votes and win the election. Lucy would decide not to run in the election when she would have won a seat on the City Council.

Each type of error has a probability of occurring. If the null hypothesis is true, the level of significance, α , is also the probability of a Type I error, while the probability of a Type II error is denoted by β .

Truth	Decision Made	Result	Associated Probability
H_0 True	Reject H_0	Type I Error	α
	Do Not Reject H_0	Correct Decision	$1 - \alpha$
H_a True	Reject H_0	Correct Decision	$1 - \beta = \text{power}$
	Do Not Reject H_0	Type II Error	β

Another important component of a hypothesis test is its **power**. The **power** of a test measures its ability to detect an alternative hypothesis when it is true. Power of a particular test is calculated as the probability that the test will reject H_0 when the alternative hypothesis is true.

For Lucy's test, the power is calculated as the probability that we conclude Lucy will receive more than $\frac{1}{2}$ the votes in the election when the truth is that she actually will win the election. Since we just learned that β is the probability that we did NOT reject H_0 when the H_a is true, we can see that power is represented by $1 - \beta$.

Formula Card

Corresponding details and formulas as they appear on the Stats 250 formula card are shown at the right.

Population Proportion	
Parameter	p
Statistic	\hat{p}
Standard Error	$s.e.(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
Confidence Interval	$\hat{p} \pm z^* s.e.(\hat{p})$
Conservative Confidence Interval	$\hat{p} \pm \frac{z^*}{2\sqrt{n}}$
Large-Sample z-Test	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
Sample Size	$n = \left(\frac{z^*}{2m} \right)^2$

Warm-Up: Stating the Hypotheses and Defining the Parameter of Interest

For each example, fill in the hypotheses and define the parameter of interest.

1. The Detroit Tigers advertising team believes that about 70% of Ann Arbor residents will attend a Tigers game this season. Suppose the general manager wants to cut back on advertising in the area and speculates that Ann Arbor residents are less likely to attend a Detroit Tigers game this season than previously assumed.

Let the parameter ____ represent

H_0 : _____ H_a : _____

2. The University of Michigan is interested in examining the proportion of students who graduate in four years. Suppose that a researcher speculates that the University has a majority of students who graduate four years after they enroll.

Let the parameter ____ represent

H_0 : _____ H_a : _____

ILP: Hypothesis Testing for a Population Proportion

In this ILP, we are going to use a sample proportion to test a theory about the value of the population proportion. Again, we assume the lab section is a representative random sample of the UM student body.

1. Determine a question to investigate by filling in the following question.

“What proportion of UM students ...
_____?”

2. Once the question has been determined, the instructor will pose the question to the class using clickers. Record the results below and provide an estimate of the population proportion of all UM students that _____ (that is, report the sample proportion).

# of students responding:	
# of student who responded “yes”:	
Sample proportion of student who responded “yes” : (symbol and value)	

3. The estimate in part (2) is a sample proportion. About how far away from the population proportion would you expect such estimates to be, on average? (i.e. report the standard error of the sample proportion).

4. We wish to use the data collected to assess:
"Do a ... (select one) **minority** **majority** of UM students _____?"
Express this as an appropriate null and alternative hypothesis:

H_0 : _____ H_a : _____

where p represents: _____

5. A large sample Z test can be performed using a normal approximation for the binomial if the sample size is large enough. Provide the checks necessary to see if this approximation may be made.
6. Assume the sample is large enough and calculate the corresponding z test statistic value. Show your work.

What is the distribution of the test statistic under the null hypothesis?

7. Use your test statistic to find the corresponding p -value. First make a quick sketch by hand and use Table A.1 to find your p -value; then try the new `pval()` script in R to check your answer.

What would the p -value look like if the alternative hypothesis had been that the population proportion did not equal 0.5?

8. What is your decision at the 5% significance level? **Reject H_0** **Fail to Reject H_0**

Also write out your real world conclusion in the context of the problem.

Cool-Down: Errors and Power Computation

Maize and Blue Marbles: Cory is going to play the game 'What is in the Box?' He will win season tickets to the Red Wings if he can correctly identify the contents of a box (which he cannot see). There are two possibilities for the box contents:

H_0 : The box contains eight red marbles and four white marbles.

H_a : The box contains two red marbles and ten white marbles.

Cory will select one marble from the box without looking. He must make his decision based on the color of that one selected marble. He has picked the following decision rule to use: **Reject H_0 if the observed color of the selected marble is WHITE.**

It may help to make a quick visual picture of these two boxes.

- a. For this situation, what is the probability of incorrectly rejecting H_0 ?

2/12 4/12 10/12

- b. Which of the following terms is appropriate for the probability described in part (a)?

Type I error Type II error Power

- c. For this situation, what is the probability of incorrectly failing to reject H_0 ?

2/12 4/12 10/12

- d. Which of the following terms is appropriate for the probability described in part (c)?

Type I error Type II error Power

- e. For this situation, what is the power?

2/12 4/12 10/12
