

8.7: Hypothesis Testing (1 of 5)

Learning Objectives

- When testing a claim, distinguish among situations involving one population mean, one population proportion, two population means, or two population proportions.
- Given a claim about a population, determine null and alternative hypotheses.

Introduction

In inference, we use a sample to draw a conclusion about a population. Two types of inference are the focus of our work in this course:

- Estimate a population parameter with a confidence interval.
- Test a claim about a population parameter with a hypothesis test.

We can also use samples from two populations to compare those populations. In this situation, the two types of inference focus on differences in the parameters.

- Estimate a difference in population parameters with a confidence interval.
- Test a claim about a difference in population parameters with a hypothesis test.

In “Estimating a Population Proportion,” we learned to estimate a population proportion using a confidence interval. For example, we estimated the proportion of all Tallahassee Community College students who are female and the proportion of all American adults who used the Internet to obtain medical information in the previous month. We will revisit confidence intervals in future modules.

Now we look more carefully at how to test a claim with a hypothesis test. Statistical investigations begin with research questions. We begin our discussion of hypothesis tests with research questions that require us to test a claim. Later we look at how a claim becomes a hypothesis.

Example

Research Questions about Testing Claims



Let’s revisit some of the research questions from examples in the module *Types of Statistical Studies and Producing Data* that involve testing a claim.

Is the average course load for community college students less than 12 semester hours? This question contains a claim about a population mean. The question contains information about the population, the variable, and the parameter. The population is all community college students. The variable is *course load in semester hours*. It is quantitative, so the parameter is a mean. The claim is, “The mean course load for all community college students is less than 12 semester hours.”

Do the majority of community college students qualify for federal student loans? This question contains a claim about a population proportion and information about the population, the variable, and the parameter. The population is all community college students. The variable is *Qualify for federal student loan* (yes or no). It is categorical, so the parameter is a proportion. The claim is, “The proportion of community college students who qualify is greater than 0.5” (a majority means more than half, or 0.5).

In community colleges, do female students and male students have different mean GPAs? This question contains a claim that compares two population means. Again, we see information about the populations, the variable, and the parameters. The two populations are female community college students and male community college students. The variable is *GPA*. It is quantitative, so the parameters are means. The claim is, “The mean GPA for female community college students is different from the mean GPA for male community college students.” Notice that the claim compares the two population means, but there is no claim about the numeric value of either mean.

Are college athletes more likely than nonathletes to receive academic advising? This question contains a claim that compares two population proportions: college athletes and college students who are not athletes. The variable is *Receive academic advising* (yes or no). The variable is categorical, so the parameters are proportions. The claim is, “The proportion of all college athletes who receive academic advising is greater than the proportion of all nonathletes in college who receive academic advising.” Notice that the claim compares two population proportions, but there is no claim about the numeric value of either proportion.

In the case of testing a claim about a single population parameter, we compare it to a numeric value. In the case of testing a claim about two population parameters, we compare them to each other.

Try It

Identify the type of claim in each research question below.

<https://assessments.lumenlearning.co...sessments/3960>

<https://assessments.lumenlearning.co...sessments/3961>

<https://assessments.lumenlearning.co...sessments/3962>

<https://assessments.lumenlearning.co...sessments/3963>

Next Steps: Forming Hypotheses

We already know that in inference we use a sample to draw a conclusion about a population. If the research question contains a claim about the population, we translate the claim into two related hypotheses.

The **null hypothesis** is a hypothesis about the value of the parameter. The null hypothesis relates to our work in *Linking Probability to Statistical Inference* where we drew a conclusion about a population parameter on the basis of the sampling distribution. We started with an assumption about the value of the parameter, then used a simulation to simulate the selection of random samples from a population with this parameter value. Or we used the parameter value in a mathematical model to describe the center and spread of the sampling distribution. *The null hypothesis gives the value of the parameter that we will use to create the sampling distribution.* In this way, the null hypothesis states what we assume to be true about the population.

The **alternative hypothesis** usually reflects the claim in the research question about the value of the parameter. The alternative hypothesis says the parameter is “greater than” or “less than” or “not equal to” the value we assume to true in the null hypothesis.

Example

Stating Hypotheses

Here are the hypotheses for the research questions from the previous example. The null hypothesis is abbreviated H_0 . The alternative hypothesis is abbreviated H_a .

Is the average course load for community college students less than 12 semester hours?

H_0 : The mean course load for community college students is equal to 12 semester hours.

H_a : The mean course load for community college students is less than 12 semester hours.

Do the majority of community college students qualify for federal student loans?

H_0 : The proportion of community college students who qualify for federal student loans is 0.5.

H_a : The proportion of community college students who qualify for federal student loans is greater than 0.5.

When the research question contains a claim that compares two populations, the null hypothesis states that the parameters are equal. We will see in Modules 9 and 10 that we translate the null hypothesis into a statement about “no difference” in parameter values. We revisit this idea in more depth later.

In community colleges, do female students and male students have different mean GPAs?

H_0 : In community colleges, female and male students have the same mean GPAs.

H_a : In community colleges, female and male students have different mean GPAs.

Are college athletes more likely than nonathletes to receive academic advising?

H_0 : In colleges, the proportion of athletes who receive academic advising is equal to the proportion of nonathletes who receive academic advising.

H_a : In colleges, the proportion of athletes who receive academic advising is greater than the proportion of nonathletes who receive academic advising.

Comment

Here are some general observations about null and alternative hypotheses.

- The hypotheses are competing claims about the parameter or about the comparison of parameters.
- Both hypotheses are statements about the same population parameter or same two population parameters.
- The null hypothesis contains an equal sign.
- The alternative hypothesis is always an inequality statement. It contains a “less than” or a “greater than” or a “not equal to” symbol.
- In a statistical investigation, we determine the research question, and thus the hypotheses, before we collect data.

The process of forming hypotheses, collecting data, and using the data to draw a conclusion about the hypotheses is called **hypothesis testing**.

Try It

<https://assessments.lumenlearning.co...sessments/3596>

<https://assessments.lumenlearning.co...sessments/3597>

Try It

<https://assessments.lumenlearning.co...sessments/3598>

<https://assessments.lumenlearning.co...sessments/3599>

Contributors and Attributions

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** [CC BY: Attribution](#)

This page titled [8.7: Hypothesis Testing \(1 of 5\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Lumen Learning](#).